

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:30:36 ; Search time 55 Seconds
(without alignments)
498.311 Million cell updates/sec

Title: US-10-676-296-2_COPY_1_97

Perfect score: 506

Sequence: 1 MLQTPQLFKNLSDTRLG.....PSWAAQIEQAWLSQDDVQ 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseqp_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	117	6	ABU63440 B. coli A
2	431	85.2	109	6	ABU31952 Protein e
3	414	81.8	117	6	ABU15462 Protein e
4	414	81.8	973	4	ABG24942 Novel hum
5	399	78.9	113	6	ABU49785 Protein e
6	349	69.0	106	6	ABU27992 Protein e
7	327	64.6	111	6	ABU27449 Protein e
8	312	61.7	576	4	ABG18368 Novel hum
9	303	59.9	111	6	ABU41107 Protein e
10	282	55.7	115	6	ABU40241 Protein e
11	266	52.6	116	6	ABU38444 Protein e
12	245	48.4	110	6	ADA36925 Acinetoba
13	228	45.1	113	6	ABU49297 Protein e
14	227	44.9	109	6	ABU16869 Protein e
15	227	44.9	116	6	ADA35167 Acinetoba
16	151	29.8	126	6	ABU36828 Protein e
17	151	29.8	126	6	ABU34403 Protein e
18	136	26.9	111	5	ABU48872 Listeria
19	136	26.9	111	6	ABU33098 Protein e
20	135	26.7	101	6	ABU17506 Protein e
21	133	26.4	107	6	ABU19574 Protein e
22	130.5	25.8	107	6	ABU24703 Protein e
23	128.5	25.4	123	4	AAB76789 Corynebac
24	128.5	25.4	123	4	AAB79020 C. glutam
25	128.5	25.4	129	4	AAG92802 C glutami

ALIGNMENTS

RESULT 1

ABU63440
ID ABU63440 standard; protein; 117 AA.
XX
AC ABU63440;
XX
DT 10-OCT-2003 (first entry)
XX
DE E. coli ArsrR protein.
XX
KW ArsrR; arsenic resistance operon; biosensor; codon optimisation; arsenic.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Region 1..97
FT /note= "Thus region is specifically claimed in claim 33"
XX
PN US2003096275-A1.
XX
PD 22-MAY-2003.
XX
PF 15-AUG-2002; 2002US-00222952.
XX
PA 20-AUG-2001; 2001US-0313714P.
XX
PI (LAIN/) LAING L G.
XX
PI Laing LG;
XX
XX WPI; 2003-576876/54.
XX N-PSDB; ACD28583.
XX
XX New system comprising isolated protein and nucleic acid, and a detection system that indicates a change in binding of the protein to the nucleic acid in the presence of the analyte, useful for detecting the presence of analyte in a sample.
XX
XX Claim 32; Page 16; 36pp; English.

The invention relates to a new system (biosensor) for detecting the presence of analyte in a sample comprising: (1) an isolated protein that specifically binds the analyte; (2) an isolated nucleic acid containing a specific binding sequence that is bound specifically by the protein; and (3) a detection system that indicates a change in binding of the protein to the nucleic acid in the presence of the analyte. Also included are a biosensor device for detecting the presence of an analyte in a sample, detecting the presence of an analyte in a sample, and an ArsrR (encoded by

part of the arsenic resistance operon of E. coli) protein comprising an amino acid sequence that is at least 90% identical to amino acids 1-97 of the ArgR sequence appearing as ABU63440 binding to a nucleic acid sequence comprising oligonucleotides appearing as ACD28584-ACD28591. The system is useful for detecting the presence of analyte in a sample. The present sequence is the E. coli ArgR protein (encoded by a codon optimised DNA) used in the biosensor of the invention to detect arsenic

Sequence 117 AA;
 Query Match 100.0%; Score 506; DB 6; Length 117;
 Best Local Similarity 100.0%; Pred. No. 5.8e-54; Indels 0; Gaps 0;
 Matches 97; Conservative 0; Mismatches 0; Mismatches 0; Gaps 0;

QY 1 MLQTLPLQFNLSDETRIGVILLREMGELVCVCDLQALDQSQPKISRHLMLRESGIL 60
 DB 1 MLQTLPLQFNLSDETRIGVILLREMGELVCVCDLQALDQSQPKISRHLMLRESGIL 60

QY 61 LDRKQGWVHYRLSPHIPSWAAQIIIEQAWLSQQDDVQ 97
 DB 61 LDRKQGWVHYRLSPHIPSWAAQIIIEQAWLSQQDDVQ 97

RESULT 2
 ABU31952
 ID ABU31952 standard; protein; 109 AA.
 XX AC ABU31952;
 XX DT 19-JUN-2003 (first entry)
 XX DE Protein encoded by Prokaryotic essential gene #17479.
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Klebsiella pneumoniae.
 XX FN WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX DR WPI; 2003-029926/02.
 XX DR N-PSDB; ACA35822.
 XX PT New antisense nucleic acids, useful for identifying proteins or screening
 XX PT for homologous nucleic acids required for cellular proliferation to
 XX PT isolate candidate molecules for rational drug discovery programs.
 XX PS Claim 25; SEQ ID NO 59876; 1765pp; English.
 XX CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for

proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than S. aureus, S. typhimurium, K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of the target prokaryotic essential genes. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 109 AA;
 Query Match 85.2%; Score 431; DB 6; Length 109;
 Best Local Similarity 83.3%; Pred. No. 8.5e-45;
 Matches 80; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 2 LQTLPLQFNLSDETRIGVILLREMGELVCVCDLQALDQSQPKISRHLMLRESGIL 61
 DB 1 MSLLPLQFNLSDETRIGVILLREMGELVCVCDLQALDQSQPKISRHLMLRESGIL 60

QY 62 DRKQGWVHYRLSPHIPSWAAQIIIEQAWLSQQDDVQ 97
 DB 61 DRKQGWVHYRLSPHIPSWAAQIIIEQAWLSQQDDVQ 96

RESULT 3
 ABU15462
 ID ABU15462 standard; protein; 117 AA.
 XX AC ABU15462;
 XX DT 19-JUN-2003 (first entry)
 XX DE Protein encoded by Prokaryotic essential gene #989.
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Escherichia coli.
 XX FN WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948993.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX DR WPI; 2003-029926/02.
 XX DR N-PSDB; ACA19332.
 XX PT New antisense nucleic acids, useful for identifying proteins or screening
 XX PT for homologous nucleic acids required for cellular proliferation to

PR isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 43386; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 117 AA;

Query Match 81.8%; Score 414; DB 6; Length 117;

Best Local Similarity 80.9%; Pred. No. 1.1e-42;

Matches 76; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY 4 LTPQLFKNLSDTRIGIVLLFEMGELCVCDLCMALDQSQPKISRHLMLRESGLILDR 63

Db 4 LTPQLFKNLSDTRIGIVLLFEMGELCVCDLCMALDQSQPKISRHLMLRESGLILDR 63

QY 64 KQKQWVHYRLSPHPSPAQAIIIEQAWLSQDDVQ 97

Db 64 KQKQWVHYRLSPHPSPAQAIIIEQAWLSQDDVQ 97

RESULT 4

ABG24942

ID ABG24942 standard; protein; 973 AA.

XX AC ABG24942;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #24933.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS89129.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.

PS Claim 20; SEQ ID NO 55301; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activities. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 973 AA;

Query Match 81.8%; Score 414; DB 4; Length 973;

Best Local Similarity 80.9%; Pred. No. 1.7e-41;

Matches 76; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY 4 LTPQLFKNLSDTRIGIVLLFEMGELCVCDLCMALDQSQPKISRHLMLRESGLILDR 63

Db 4 LTPQLFKNLSDTRIGIVLLFEMGELCVCDLCMALDQSQPKISRHLMLRESGLILDR 63

QY 64 KQKQWVHYRLSPHPSPAQAIIIEQAWLSQDDVQ 97

Db 64 KQKQWVHYRLSPHPSPAQAIIIEQAWLSQDDVQ 97

RESULT 5

ABU49785

ID ABU49785 standard; protein; 113 AA.

XX AC ABU49785;

XX 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #35312.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Versinia pestis.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.
 06-SEP-2001; 2001US-00948993.
 25-OCT-2001; 2001US-0342923P.
 08-FEB-2002; 2002US-00072851.
 06-MAR-2002; 2002US-0362699P.
 (ELIT-) ELITRA PHARM INC.
 Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 WPI; 2003-029926/02.
 N-PSDB; ACA53655.
 New antisense nucleic acids, useful for identifying proteins or screening
 for homologous nucleic acids required for cellular proliferation to
 isolate candidate molecules for rational drug discovery programs.
 Claim 25; SEQ ID NO 77709; 1766pp; English.
 The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway; (8)
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
 the target prokaryotic essential genes. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

Query Match 78.9%; Score 399; DB 6; Length 113;
 Best Local Similarity 76.0%; Pred. No. 7.6e-41;
 Matches 73; Conservative 13; Mismatches 10; Indels 0; Gaps 0;
 Sequence 113 AA;

1 MQLTFLQKFNLSDETRIGVILLREMGELCVCLDMALDQSQPKISRHLAMRESGILL 60
 1 MTLTFLQKFNLSDETRINILLKASGELCVCLCHRLNEAQPISRHLAMRESGILL 60
 61 LDKQKQWVHYRLSPHPWAQIIEQAWLSQDDV 96
 61 LDRRAGKQWVHYRLSPHPWAQIIEQAWLSQDDV 96

RESULT 6
 ABU27992
 ID ABU27992 standard; protein; 106 AA.
 XX
 AC ABU27992;
 XX

19-JUN-2003 (first entry)
 Protein encoded by Prokaryotic essential gene #13519.
 Antisense; prokaryotic essential gene; cell proliferation; drug design.
 Enterobacter cloacae.
 WO200277183-A2.
 03-OCT-2002.
 21-MAR-2002; 2002WO-US009107.
 21-MAR-2001; 2001US-00815242.
 06-SEP-2001; 2001US-00948993.
 25-OCT-2001; 2001US-0342923P.
 08-FEB-2002; 2002US-00072851.
 06-MAR-2002; 2002US-0362699P.
 (ELIT-) ELITRA PHARM INC.
 Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 WPI; 2003-029926/02.
 N-PSDB; ACA31862.
 New antisense nucleic acids, useful for identifying proteins or screening
 for homologous nucleic acids required for cellular proliferation to
 isolate candidate molecules for rational drug discovery programs.
 Claim 25; SEQ ID NO 55916; 1766pp; English.
 The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway; (8)
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
 the target prokaryotic essential genes. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

Query Match 69.0%; Score 349; DB 6; Length 106;
 Best Local Similarity 69.8%; Pred. No. 9.6e-35;
 Matches 60; Conservative 17; Mismatches 9; Indels 0; Gaps 0;
 Sequence 106 AA;

1 MQLTFLQKFNLSDETRIGVILLREMGELCVCLDMALDQSQPKISRHLAMRESGILL 63
 1 MTLTFLQKFNLSDETRINILLKASGELCVCLCHRLNEAQPISRHLAMRESGILL 63
 61 LDKQKQWVHYRLSPHPWAQIIEQAWLSQDDV 96
 61 LDRRAGKQWVHYRLSPHPWAQIIEQAWLSQDDV 96

RESULT 6
 ABU27992
 ID ABU27992 standard; protein; 106 AA.
 XX
 AC ABU27992;
 XX

Db 2 LHPLOFKTSLDETRLAIVMLREAEALCVCLVATAESQPKVSRHALLRESGLVIDR 61
 QY 64 KQGWVHYRLSPHPSWAAQIIQAW 89
 Db 62 REGKWVHYRLSPNMDAWAAVVIDNSW 87

RESULT 7
 ABU27449
 ID ABU27449 standard; protein; 111 AA.
 XX
 AC ABU27449;
 XX
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #12976.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Enterobacter cloacae.
 XX
 WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 WPI; 2003-029926/02.
 DR N-PSDB; ACA31319.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 55373; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 111 AA;
 Query Match 64.6%; Score 327; DB 6; Length 111;
 Best Local Similarity 70.4%; Pred. No. 5, 1e-32;
 Matches 57; Conservative 14; Mismatches 10; Indels 0; Gaps 0;
 QY 9 LFKNLSDETRLGIVLLREMGELCVCDLQMALDQSQPKISRHLAMRESGILLDRKQGW 68
 Db 1 LFKILSDETRLAIVMLRESGELCVCDICATSESQPKISRHMAILREAGLVLDREGKW 60
 QY 69 VHYRLSPHPSWAAQIIQAW 89
 Db 61 IHXRLSPHIPAWAETIMTSW 81

RESULT 8
 ABG18368
 ID ABG18368 standard; protein; 576 AA.
 XX
 AC ABG18368;
 XX
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #18359.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 WPI; 2001-639362/73.
 DR N-PSDB; AAS82555.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 48727; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 576 AA;

Query Match 61.7%; Score 312; DB 4; Length 576;
Best Local Similarity 82.1%; Pred. No. 2.9e-29;
Matches 55; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 31 LCVCDLCMALDQSPKISRHLAMRESGILLDRKQGWVHYRLSPHPSWAAQIIQAWL 90
Db 72 LCVCDLCTALDQSPKISRHLALRESGLLDRLKQGWVHYRLSPHPSWAAKIIDEAR 131
QY 91 SQQDDVQ 97
Db 132 CEQEKVQ 138

RESULT 9
ABU41107
ID ABU41107 standard; protein; 111 AA.
XX
AC ABU41107;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #26634.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Proteus sp.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
FA (ELIT-) ELITRA PHARM INC.
XX
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA44977.
XX
New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 69031; 1766pp; English.
XX
The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 111 AA;

Query Match 59.9%; Score 303; DB 6; Length 111;
Best Local Similarity 63.4%; Pred. No. 4.5e-29;
Matches 59; Conservative 13; Mismatches 21; Indels 0; Gaps 0;

QY 4 LTPQLFKMLSDETRLGIVLLREMGELCVCDLCMALDQSPKISRHLAMRESGILLDR 63
Db 1 MKELQLFKLIGDQTRLDIVLLKASGELCVCDIYTNLSQPKTSRLAMRESGLLDS 60
QY 64 KQGWVHYRLSPHPSWAAQIIQAWLSQDDV 96
Db 61 KQGWVHYRLSPVLLPWVKSIIIDITYTEKNRV 93

RESULT 10
ABU40241
ID ABU40241 standard; protein; 115 AA.
XX
AC ABU40241;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #25768.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Pseudomonas putida.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
FA (ELIT-) ELITRA PHARM INC.
XX
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA44111.
XX
New antisense nucleic acids, useful for identifying proteins or screening

for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 68165; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway; (8) required for proliferation, or that inhibits cellular proliferation; (9) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 115 AA;

Query Match 55.7%; Score 282; DB 6; Length 115;

Best Local Similarity 54.3%; Pred. No. 1.8e-26; Matches 51; Conservative 19; Mismatches 18; Indels 6; Gaps 1;

4 LPLQLFKNLSDTRLGIVLLREMGELCVCDLGMALDQSQPKISRHLAMLRSGILLDR 63
2 ITPDVKLSLSDTRARATLLASLGELCVCELMCALNDSPKISRHLAQLRSGMLDR 61

64 KQKWHYRLSPHPSWAAQIIE-----QAMLS 91

62 RQGWVYRLNPLPSWVHEMLQVTLQANSQWLA 95

RESULT 11

ABU38444

ID ABU38444 standard; protein; 116 AA.

XX ABU38444;

AC ABU38444;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #23971.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Pseudomonas aeruginosa.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX

XX

XX

XX

XX

XX

XX

XX

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

PR

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA42314.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX for homologous nucleic acids required for cellular proliferation to

XX isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 66368; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

XX of the nucleic acid inhibits proliferation of a cell. Also included are:

XX (1) a vector comprising a promoter operably linked to the nucleic acid

XX encoding a polypeptide whose expression is inhibited by the antisense

XX nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX polypeptide or its fragment whose expression is inhibited by the

XX antisense nucleic acid; (4) an antibody capable of specifically binding

XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX proliferation or the activity of a gene in an operon required for

XX proliferation; (7) identifying a compound that influences the activity of

XX the gene product or that has an activity against a biological pathway

XX required for proliferation, or that inhibits cellular proliferation; (8)

XX identifying a gene required for cellular proliferation or the biological

XX pathway in which a proliferation-required gene or its gene product lies

XX or a gene on which the test compound that inhibits proliferation of an

XX organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX compound's activity; (11) a culture comprising strains in which the gene

XX product is overexpressed or underexpressed; (12) determining the extent

XX to which each of the strains is present in a culture or collection of

XX strains; or (13) identifying the target of a compound that inhibits the

XX proliferation of an organism. The antisense nucleic acids are useful for

XX identifying proteins or screening for homologous nucleic acids required

XX for cellular proliferation to isolate candidate molecules for rational

XX drug discovery programs, or for screening homologous nucleic acids

XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

XX the target prokaryotic essential genes. Note: The sequence data for this

XX patent did not form part of the printed specification, but was obtained

XX in electronic format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 116 AA;

Query Match 52.6%; Score 266; DB 6; Length 116;

Best Local Similarity 51.6%; Pred. No. 1.7e-24;

Matches 48; Conservative 22; Mismatches 23; Indels 0; Gaps 0;

5 TPLQLFKNLSDTRLGIVLLREMGELCVCDLGMALDQSQPKISRHLAMLRSGILLDR 64

3 SPAEYFKCLADETRVATLLIVDQGLCVCELMCALNDSPKISRHLAQLRSGILLDR 62

65 QKQWHYRLSPHPSWAAQIIEQAMLSQDDVQ 97

63 QGQWVYRLNPLPSWVHEMLQVTLQANSQWLA 95

RESULT 12

ADA36925

ID ADA36925 standard; protein; 110 AA.

XX ADA36925;

XX ADA36925;

XX 20-NOV-2003 (first entry)

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

21-MAR-2002; 2002MO-US009107.
21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
(ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
WPI; 2003-029926/02.
N-PSDB; ACA53167.
New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
Claim 25; SEQ ID NO 77221; 1766pp; English.
The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

21-MAR-2002; 2002MO-US009107.
21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993.
25-OCT-2001; 2001US-0342923P.
08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-0362699P.
(ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
WPI; 2003-029926/02.
N-PSDB; ACA53167.
New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.
Claim 25; SEQ ID NO 77221; 1766pp; English.
The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
(1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for
proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of
strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

AC ABU16869;
XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #2396.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Acinetobacter baumannii.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA20739.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 4793; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 109 AA;
XX Query Match 44.9%; Score 227; DB 6; Length 109;
XX Best Local Similarity 49.4%; Pred. No. 9.4e-20;
XX Matches 41; Conservative 18; Mismatches 24; Indels 0; Gaps 0;

QY 4 LTPQLQFKNLSDETRIGIVLLIREMGECLVCDCWALDQSQPKISRHLAMLRSGILDR 63
DB 2 INQVDFEKLSDQTRINILKLVLNKQNCVCELTEQLSELSQPKISRHLALLRTHGVILDE 61
QY 64 KQKQWVHYRLSPHIPSWAAQIIIE 86
DB 62 RKGQWVYYSINPDLFPWALDILK 84

RESULT 15

ADA35167
ID ADA35167 standard; protein; 116 AA.

XX AC ADA35167;

XX DT 20-NOV-2003 (first entry)

XX DE Acinetobacter baumannii protein #2328.

XX KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
XX plant biocontrol agent.

XX OS Acinetobacter baumannii.

XX FN US6562958-B1.

XX PD 13-MAY-2003.

XX PF 04-JUN-1999; 99US-00328352.

XX PR 09-JUN-1998; 98US-0088701P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Breton G, Bush D;

XX DR WPI; 2003-576092/54.

XX DR N-PSDB; ADA31041.

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX for diagnosing a bacterial disease, as components of antibacterial
XX vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX plants.

XX Example; SEQ ID NO 6454; 328pp; English.

XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX The A. baumannii nucleic acids and polypeptides are useful as reagents
XX for diagnosing a bacterial disease, as components of antibacterial
XX vaccines, as targets for antibacterial drugs, to detect the presence of
XX A. baumannii and other Acinetobacter species in a sample, in screening
XX compounds for the ability to interfere with the A. baumannii life cycle
XX or to inhibit A. baumannii infection, and as biocontrol agents for
XX plants. The present sequence represents the amino acid sequence of an A.
XX baumannii protein.

XX SQ Sequence 116 AA;

Query Match 44.9%; Score 227; DB 6; Length 116;

Best Local Similarity 49.4%; Pred. No. 1e-19;

Matches 41; Conservative 18; Mismatches 24; Indels 0; Gaps 0;

QY 4 LTPQLQFKNLSDETRIGIVLLIREMGECLVCDCWALDQSQPKISRHLAMLRSGILDR 63
DB 9 INQVDFEKLSDQTRINILKLVLNKQNCVCELTEQLSELSQPKISRHLALLRTHGVILDE 68

QY 64 KQKQWVHYRLSPHIPSWAAQIIIE 86

DB 69 RKGQWVYYSINPDLFPWALDILK 91

Search completed: June 2, 2004, 19:36:54

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:35:52 ; Search time 22 Seconds
(without alignments)
227.624 Million cell updates/sec

Title: US-10-676-296-2_COPY_1_97

Perfect score: 506
Sequence: 1 MLQTLPLQFKNLSDTRGLG.....PSWAAQIIIEQAWLSQDDVQ 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	432	85.4	112	4	US-09-489-039A-11993
2	303	59.9	119	4	US-09-543-681A-7062
3	266	52.6	142	4	US-09-252-991A-25372
4	245	48.4	110	4	US-09-328-352-8212
5	227	44.9	116	4	US-09-328-352-6454
6	127.5	25.2	105	4	US-09-107-532A-6671
7	126	24.9	127	4	US-09-134-001C-2896
8	115.5	22.8	110	4	US-09-134-001C-2984
9	115	22.7	117	4	US-09-134-001C-3997
10	115	22.7	120	4	US-09-107-532A-7083
11	85	16.8	124	4	US-09-134-001C-4197
12	83.5	16.5	121	4	US-09-543-681A-1719
13	82.5	16.3	118	3	US-09-413-814-17
14	81.5	16.1	120	4	US-09-134-000C-6033
15	75.5	14.9	348	4	US-09-252-991A-19898
16	75	14.8	99	4	US-09-489-039A-9929
17	74.5	14.7	103	4	US-09-328-352-7724
18	73.5	14.5	237	4	US-09-489-039A-7709
19	72	14.2	176	4	US-09-252-991A-24407
20	72	14.2	237	4	US-09-543-681A-4561
21	70.5	13.9	210	4	US-09-543-681A-7093
22	70	13.8	644	4	US-09-198-452A-1135
23	69	13.6	320	1	US-09-252-991A-27596
24	68.5	13.5	1620	1	US-08-542-363-2
25	68.5	13.5	1620	3	US-09-100-089-2
26	68.5	13.5	1620	4	US-09-670-827-2
27	68	13.4	257	4	US-09-252-991A-19163

28	67.5	13.3	124	4	US-09-489-039A-8144	Sequence 8144, Ap
29	67.5	13.3	539	4	US-09-252-991A-24784	Sequence 24784, A
30	67.5	13.3	666	4	US-09-228-986-68	Sequence 68, Appl
31	66	13.0	1054	4	US-09-693-542-87	Sequence 87, Appl
32	65.5	12.9	282	4	US-09-252-991A-17674	Sequence 17674, A
33	65.5	12.9	676	4	US-09-134-000C-6050	Sequence 6050, Ap
34	64.5	12.7	273	4	US-09-489-039A-9484	Sequence 9484, Ap
35	64	12.6	339	4	US-09-252-991A-27733	Sequence 27733, A
36	64	12.6	1244	3	US-08-938-291A-5	Sequence 5, Appli
37	64	12.6	1244	4	US-09-589-619-5	Sequence 2, Appli
38	63.5	12.5	610	4	US-09-455-777-2	Sequence 9576, Ap
39	63	12.5	313	4	US-09-489-039A-9576	Sequence 2, Appli
40	62.5	12.4	295	2	US-08-411-607A-2	Sequence 2, Appli
41	62.5	12.4	295	4	US-09-361-741-2	Sequence 2, Appli
42	62.5	12.4	295	4	US-09-461-418-2	Sequence 4473, Ap
43	62.5	12.4	300	4	US-09-107-532A-4473	Sequence 9, Appli
44	62.5	12.4	308	3	US-08-927-433-9	Sequence 4, Appli
45	62.5	12.4	308	3	US-08-775-882-4	

ALIGNMENTS

RESULT 1
US-09-489-039A-11993
; Sequence 11993, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11993
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11993

Query Match 85.4%; Score 432; DB 4; Length 112;
Best Local Similarity 82.5%; Pred. No. 3e-49;
Matches 80; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 1 MLQTLPLQFKNLSDTRGLGIVLLREMGELVCVCDLQMALDOSOPKISRHLAMRESGIL 60
DB 3 IMSLLPLQFKNLADETRGLGIVLLKARGELVCVCDLQALQSQPKISRHLAMRESGLL 62
QY 61 LDRKQKWHYRLSPSWAAQIIIEQAWLSQDDVQ 97
DB 63 LDRKQKWHYRLSPQMPAAQVIEQAWLSQDDVQ 99

RESULT 2
US-09-543-681A-7062
; Sequence 7062, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7062
; LENGTH: 119
; TYPE: PRT

```

; ORGANISM: Proteus mirabilis
US-09-543-681A-7062

Query Match
Best Local Similarity 59.9%; Score 303; DB 4; Length 119;
Matches 59; Conservative 13; Mismatches 21; Indels 0; Gaps 0;

QY 4 LTPLQFLKNSDETRIGIVLLREMGELCVCDLQMALDQSQPKISRHLAMLRRESGILLDR 63
Db 9 MKPLQFLKILGQDTRLDIVLLKASGELCVCDIYTAIQLNSQPKISRHLAMLRRESGILLDRS 68

QY 64 KQKQWVHYRLSHIPSWAAQIIEQAWLSQDDV 96
Db 69 KQKQWVHYRLSPVLEPWKNRIIDITYTTEKNRV 101

RESULT 3
US-09-252-991A-25372
; Sequence 25372, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25372
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25372

Query Match
Best Local Similarity 52.6%; Score 266; DB 4; Length 142;
Matches 48; Conservative 22; Mismatches 23; Indels 0; Gaps 0;

QY 5 TPLQFLKNSDETRIGIVLLREMGELCVCDLQMALDQSQPKISRHLAMLRRESGILLDRX 64
Db 29 SPAEVFKLADETRVRATLLVDQGEVCELMCALDQSQPKISRHLAQLRSAGILLDR 88

QY 65 QKQWVHYRLSHIPSWAAQIIEQAWLSQDDVQ 97
Db 89 QGQWVYRLNTPFWIHEVLTIRANGDWLQ 121

RESULT 4
US-09-328-352-8212
; Sequence 8212, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8212
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8212

Query Match
Best Local Similarity 48.4%; Score 245; DB 4; Length 110;
Matches 45; Conservative 16; Mismatches 23; Indels 0; Gaps 0;

```

```

QY 3 QLTPLQFLKNSDETRIGIVLLREMGELCVCDLQMALDQSQPKISRHLAMLRRESGILLDR 62
Db 3 KMDQVNFKLSDETRINIVLIAENNELCVCDLTKELQSQPKISRHLALLRSSGILLQD 62

QY 63 KQKQWVHYRLSHIPSWAAQIIE 86
Db 63 RRQSQWVYYSINQQLPAWCFEILD 86

RESULT 5
US-09-328-352-6454
; Sequence 6454, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6454
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6454

Query Match
Best Local Similarity 44.9%; Score 227; DB 4; Length 116;
Matches 41; Conservative 18; Mismatches 24; Indels 0; Gaps 0;

QY 4 LTPLQFLKNSDETRIGIVLLREMGELCVCDLQMALDQSQPKISRHLAMLRRESGILLDR 63
Db 9 INQVDFKLSQDTRNLKILKQNICVCELTQELQSQPKISRHLALLRTHGVLLDE 68

QY 64 KQKQWVHYRLSHIPSWAAQIIE 86
Db 69 RKQSQWVYYSINQQLPAWCFEILD 91

RESULT 6
US-09-107-532A-6671
; Sequence 6671, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6671:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...105
SEQUENCE DESCRIPTION: SEQ ID NO: 6671:
US-09-107-532A-6671

Query Match 25.2%; Score 127.5; DB 4; Length 105;
Best Local Similarity 38.2%; Pred. No. 3.5e-09;
Matches 29; Conservative 16; Mismatches 30; Indels 1; Gaps 1;
QY 8 QLFKNLSDETRGLGIVLLREMGELCVCDLGMALDQSQPKISRHLAMLRSGILLDRKQK 67
DB 13 KIFKALCDPKRLTILDYLS-GEKCAVLNIENNIQSALSVMKILCDSGIVNARQEGK 71
QY 68 WYHRLSPHPSWAAQ 83
DB 72 WTHYLSKSGSEYASK 87

RESULT 7
US-09-134-001C-2896
Sequence 2896, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 2896
LENGTH: 127
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2896

Query Match 24.9%; Score 126; DB 4; Length 127;
Best Local Similarity 39.1%; Pred. No. 7.2e-09;
Matches 27; Conservative 15; Mismatches 25; Indels 2; Gaps 1;
QY 8 QLFKNLSDETRGLGIVLLREMG--ELCVCDLGMALDQSQPKISRHLAMLRSGILLDRKQ 65
DB 42 QIFNALADKIRLKLHLSIROSNTKSLCVCDDLELLELKQSLSYHLKGLVDNIIAEXH 101
QY 66 GWVHYRLS 74
DB 102 GTWNYKIN 110

RESULT 8
US-09-134-001C-2984
Sequence 2984, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 2984
LENGTH: 110
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2984

Query Match 22.8%; Score 115.5; DB 4; Length 110;
Best Local Similarity 42.4%; Pred. No. 1.4e-07;
Matches 28; Conservative 11; Mismatches 26; Indels 1; Gaps 1;
QY 9 LFPKNLSDETRGLGIVLLREMGELCVCDLGMALDQSQPKISRHLAMLRSGILLDRKQK 68
DB 14 ILKVLSDPSRLILDLL-SCGELCACDILAYFQSQPTLSHMKILVDNELVSTRKDGTK 72
QY 69 VHYRLS 74
DB 73 RMYRLN 78

RESULT 9
US-09-134-001C-3997
Sequence 3997, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3997
LENGTH: 117
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3997

Query Match 22.7%; Score 115; DB 4; Length 117;
Best Local Similarity 36.9%; Pred. No. 1.8e-07;
Matches 24; Conservative 14; Mismatches 27; Indels 0; Gaps 0;
QY 7 LQLFKNLSDETRGLGIVLLREMGELCVCDLGMALDQSQPKISRHLAMLRSGILLDRKQ 66
DB 35 LNLKICDEKKILSLIKELCVCDSILIKMSVASTSHHRLLYKNVLDIFYKEG 94
QY 67 KWVHY 71
DB 95 KWAYY 99

RESULT 10
US-09-107-532A-7083
Sequence 7083, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:

ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7083:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...120
SEQUENCE DESCRIPTION: SEQ ID NO: 7083:
US-09-107-532A-7083
Query Match 22.7%; Score 115; DB 4; Length 120;
Best Local Similarity 39.7%; Pred. No. 1.9e-07;
Matches 27; Conservative 9; Mismatches 32; Indels 0; Gaps 0;
QY 7 LQFNLSDETRLGIVLLIREMGEICVCDLCWALDQSQPKISRHLAMLRRESGILLDRKQG 66
Db 31 LVLGKCFSDSRKIFVALETYKENCVCVDLAELTASVATTSHHLRFLKXGMAKSRDQG 90
QY 67 KWHYRLS 74
Db 91 KWIYSLA 98
RESULT 11
US-09-134-001C-4197
Sequence 4197, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4197
LENGTH: 124
TYPE: PRT

ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4197
Query Match 16.8%; Score 85; DB 4; Length 124;
Best Local Similarity 28.4%; Pred. No. 0.0017;
Matches 23; Conservative 21; Mismatches 35; Indels 2; Gaps 2;
QY 9 LFKNLSDETRLGIVLLIREMGEICVCDLCWALDQSQPKISRHLAMLRRESGILLDRKQG 69
Db 32 IFKALSDFNVRIMEFL-ENGEASVGHISHSLNMTQSNVSHQKLLKSTHLVKSRQGS 90
QY 69 VHYRLSP-HIPSWAAQIIIEQA 88
Db 91 MIYSIDDIHVSTILKQAIHHS 111
RESULT 12
US-09-543-681A-7179
Sequence 7179, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7179
LENGTH: 121
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-7179
Query Match 16.5%; Score 83.5; DB 4; Length 121;
Best Local Similarity 31.8%; Pred. No. 0.0026;
Matches 21; Conservative 15; Mismatches 29; Indels 1; Gaps 1;
QY 9 LFKNLSDETRLGIVLLIREMGEICVCDLCWALDQSQPKISRHLAMLRRESGILLDRKQG 68
Db 37 MLKTLGNGDRL-LLLCQLSQGKSVSELSLIGIRQPTLSQQLTVLRNEGIVNTRDCKR 95
QY 69 VHYRLS 74
Db 96 IFYSIA 101
RESULT 13
US-09-413-814-17
Sequence 17, Application US/09413814
Patent No. 6225064
GENERAL INFORMATION:
APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
TITLE OF INVENTION: heteropolyketide compounds
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-17

Query Match      16.3%; Score 82.5; DB 3; Length 118;
Best Local Similarity 36.1%; Pred. No. 0.0035;
Matches 22; Conservative 10; Mismatches 24; Indels 5; Gaps 2;

QY      4 LTPLOQ----LFGKNSDETRLGIVLLIREMGELCVCDLCMALDQSQPKISRHLAMLRRESGI 59
Db      1 MTPSERLDATPAALADPTERRAILARLAS-GEASVTELAKPFAMSQPALSXLKVLKVLRAGL 59

QY      60 L 60
Db      60 I 60

RESULT 14
US-09-134-000C-6033
; Sequence 6033, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6033
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6033

Query Match      16.1%; Score 81.5; DB 4; Length 120;
Best Local Similarity 33.3%; Pred. No. 0.0048;
Matches 22; Conservative 14; Mismatches 25; Indels 5; Gaps 2;

QY      13 LSDETRLGIVLLIREMGELC----VCDLCMALDQSQPKISRHLAMLRRESGILLDRKQKQW 68
Db      30 LGDEKRAIITALLE-DQACVGRVTELTQLSRPAPVSHHLKILKQAKLIDCRSEGTK 88

QY      69 VHYRLS 74
Db      89 NYYSLS 94

RESULT 15
US-09-252-991A-19898
; Sequence 19898, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19898
; LENGTH: 348
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19898

Query Match      14.9%; Score 75.5; DB 4; Length 348;
Best Local Similarity 25.3%; Pred. No. 0.12;
Matches 22; Conservative 16; Mismatches 48; Indels 1; Gaps 1;

QY      9 LFKNLSDETRLGIVLLIREMGELCVCDLCMALDQSQPKISRHLAMLRRESGILLDRKQKQW 68
Db      34 LCKAAGDSLRLNVLRALAN-DSFGVLELAQIIFAIGQSGMSHLKVLQAQAGLVATRREGNA 92

QY      69 VHYRLSHIPSWAAQIIIEQAWLSQDD 95
Db      93 IFYRRSLPLAERTGGALHAALLEEVEDE 119

Search completed: June 2, 2004, 19:39:23
Job time : 23 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:38:22 ; Search time 43 Seconds
(without alignments)
634.648 Million cell updates/sec

Title: US-10-676-296-2_COPY_1_97

Perfect score: 506

Sequence: 1 MLQTLPLQFNLSDETRIG.....PSWAAQIIIEQWLSSQDDVQ 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 281338677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	506	100.0	117	14	US-10-222-952A-2
2	431	85.2	109	12	Sequence 2, Appli
3	414	81.8	117	12	Sequence 59876, A
4	399	78.9	113	12	Sequence 43386, A
5	349	69.0	106	12	Sequence 77709, A
6	327	64.6	111	12	Sequence 55916, A
7	303	59.9	111	12	Sequence 53373, A
8	282	55.7	115	12	Sequence 69031, A
9	266	52.6	116	12	Sequence 68165, A
10	228	45.1	113	12	Sequence 66368, A
11	227	44.9	109	12	Sequence 77221, A
12	160	31.6	134	14	Sequence 44793, A
13	151	29.8	126	12	Sequence 11291, A
14	151	29.8	126	12	Sequence 63227, A
15	137.5	27.2	98	14	Sequence 64752, A
					Sequence 7873, Ap

16	136	26.9	111	12	US-10-282-122A-61022
17	135	26.7	101	12	Sequence 61022, A
18	133.5	26.4	107	12	Sequence 45430, A
19	130.5	25.8	107	12	Sequence 47498, A
20	128.5	25.4	123	12	Sequence 52627, A
21	128.5	25.4	129	9	Sequence 560, App
22	126.5	25.0	89	12	Sequence 6556, App
23	124.5	24.6	95	12	Sequence 53884, A
24	122	24.1	122	12	Sequence 56922, A
25	121.5	24.0	104	12	Sequence 52159, A
26	121.5	24.0	104	12	Sequence 70636, A
27	119.5	23.6	116	12	Sequence 243218, A
28	117	23.1	115	12	Sequence 50310, A
29	116	22.9	105	12	Sequence 47297, A
30	115.5	22.8	134	12	Sequence 71919, A
31	113	22.3	112	12	Sequence 62126, A
32	111	21.9	121	12	Sequence 74810, A
33	110.5	21.8	104	12	Sequence 53184, A
34	110.5	21.8	115	14	Sequence 70374, A
35	108.5	21.4	108	12	Sequence 9869, Ap
36	108.5	21.4	108	12	Sequence 63757, A
37	103	20.4	59	12	Sequence 77122, A
38	97	19.2	112	12	Sequence 54613, A
39	96.5	19.1	125	14	Sequence 51467, A
40	95.5	18.9	93	12	Sequence 14572, A
41	94.5	18.7	124	12	Sequence 52675, A
42	90	17.8	104	9	Sequence 49733, A
43	89	17.6	119	12	Sequence 4472, Ap
44	89	17.6	374	14	Sequence 62652, A
45	87.5	17.3	99	12	Sequence 12217, A
					Sequence 61393, A

ALIGNMENTS

RESULT 1
US-10-222-952A-2
; Sequence 2, Application US/10222952A
; Publication No. US20030096275A1
; GENERAL INFORMATION:
; APPLICANT: Regenesis
; APPLICANT: Laing, Lance
; TITLE OF INVENTION: BIOSENSOR FOR SMALL MOLECULE ANALYTES
; FILE REFERENCE: 4107/IL443-US1
; CURRENT APPLICATION NUMBER: US/10/222,952A
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/313,714
; PRIOR FILING DATE: 2001-08-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Escherichia coli
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / CAA 34168
; DATABASE ENTRY DATE: 1994-09-07
; RELEVANT RESIDUES: (1)..(117)
US-10-222-952A-2

Query Match 100.0%; Score 506; DB 14; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.4e-54;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLQTLPLQFNLSDETRIGVLLIRENGELCVCDLQALDOSQPKISRHLAMLRSGIL	60
DB	1	MLQTLPLQFNLSDETRIGVLLIRENGELCVCDLQALDOSQPKISRHLAMLRSGIL	60
QY	61	LDRKQKWWHYRLSPHPSWAAQIIIEQWLSSQDDVQ	97
DB	61	LDRKQKWWHYRLSPHPSWAAQIIIEQWLSSQDDVQ	97


```

APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43386
LENGTH: 117
TYPE: PRT
ORGANISM: Escherichia coli
US-10-282-122A-43386

Query Match      81.8%; Score 414; DB 12; Length 117;
Best Local Similarity 80.9%; Pred. No. 5.4e-43;
Matches 76; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY      4  L L P L Q L K N L S D E T R I G I V L L R E M G E L C V C D L C M A L D O S Q P K I S R H L A M L R E S G L L D R 63
Db      4  L L P I Q L F K L A D E T R I G I V L L S E L G E L C V C D L C T A L D O S Q P K I S R H L A L L R E S G L L D R 63

QY      64  K Q K W H Y L S P H I S P W A A Q I I E Q A M L S O O D D V Q 97
Db      64  K Q K W H Y L S P H I P A W A A K I I D E A W R C S Q E K V Q 97

RESULT 4
US-10-282-122A-77709
; Sequence 77709, Application US/10282122A
; Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A

```

```
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77709
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Yersinia pestis
US-10-282-122A-77709

Query Match 78.9%; Score 399; DB 12; Length 113;
Best Local Similarity 76.0%; Pred. No. 3.7e-41;
Matches 73; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLQLTPLQKFNLSDETRIGVILLREMGELCVCDLQWALDQSPKISRHLAMLRSGILL 60
Db 1 MTTLTPLQKFNLSDETRINILLKASGELCVCELCHRLNEAQPKISRHLAMLRSGILL 60

QY 61 LDRKQKWHYRLSPHPSWAAQITEQWLSQDDV 96
Db 61 LDRRAGKWHYRLSPHPSWAAQITEQWLSQDDV 96

RESULT 5
US-10-282-122A-55916
; Sequence 55916, Application US/10282122A
; Publication NO. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangau
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
```

```
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55916
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55916

Query Match 69.0%; Score 349; DB 12; Length 106;
Best Local Similarity 69.8%; Pred. No. 5e-35;
Matches 60; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 4 LTPQLKFNLSDETRIGVILLREMGELCVCDLQWALDQSPKISRHLAMLRSGILLDR 63
Db 2 LHPQLFKTSLDETRIAIVMLLREAAELCVCDLQWALDQSPKISRHLAMLRSGILLDR 61

QY 64 KQKQKWHYRLSPHPSWAAQITEQW 89
Db 62 REGKQKWHYRLSPHPSWAAQITEQW 87

RESULT 6
US-10-282-122A-55373
; Sequence 55373, Application US/10282122A
; Publication NO. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangau
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
```

```

; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55373
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (63)..(63)
; OTHER INFORMATION: X-any amino acid
;
US-10-282-122A-55373

Query Match          64.6%; Score 327; DB 12; Length 111;
Best Local Similarity 70.4%; Pred. No. 2.7e-32;
Matches 57; Conservative 14; Mismatches 10; Indels 0; Gaps 0;

QY 9 LFNKLSDETLGIVLLREMGELCVCDLCLMALDOSQPKISRHLAMLRSGILLDRKQGW 68
Db 1 LFNKLSDETLGIVLLREMGELCVCDLCLMALDOSQPKISRHLAMLRSGILLDRKQGW 60

QY 69 VHYRLSPHSPWAAQIEQAW 89
Db 61 IHRLSPHPAWAETIMTSW 81

RESULT 7
US-10-282-122A-69031
; Sequence 69031, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69031
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Pseudomonas putida
;
US-10-282-122A-69165

Query Match          55.7%; Score 282; DB 12; Length 115;
Best Local Similarity 54.3%; Pred. No. 1e-26;
Matches 51; Conservative 19; Mismatches 18; Indels 6; Gaps 1;

QY 4 LTPQLQFNLSDETLGIVLLREMGELCVCDLCLMALDOSQPKISRHLAMLRSGILLDR 63
Db 1 MKPLQLFKILGQTRLDIVLLKASGELCVCDIYTAIINLSQPKTSRHLAMLRSGILLDR 60

QY 64 KQKQWVHYRLSPHSPWAAQIEQAWLSQDDV 96
Db 61 KQKQWVHYRLSPVLLPFWVKSIIDIITYTTEKNRV 93

RESULT 8
US-10-282-122A-68165
; Sequence 68165, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68165
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Pseudomonas putida
;
US-10-282-122A-68165

Query Match          55.7%; Score 282; DB 12; Length 115;
Best Local Similarity 54.3%; Pred. No. 1e-26;
Matches 51; Conservative 19; Mismatches 18; Indels 6; Gaps 1;

QY 4 LTPQLQFNLSDETLGIVLLREMGELCVCDLCLMALDOSQPKISRHLAMLRSGILLDR 63
Db 1 MKPLQLFKILGQTRLDIVLLKASGELCVCDIYTAIINLSQPKTSRHLAMLRSGILLDR 60
```

```

; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55373
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (63)..(63)
; OTHER INFORMATION: X-any amino acid
;
US-10-282-122A-55373

Query Match          64.6%; Score 327; DB 12; Length 111;
Best Local Similarity 70.4%; Pred. No. 2.7e-32;
Matches 57; Conservative 14; Mismatches 10; Indels 0; Gaps 0;

QY 9 LFNKLSDETLGIVLLREMGELCVCDLCLMALDOSQPKISRHLAMLRSGILLDRKQGW 68
Db 1 LFNKLSDETLGIVLLREMGELCVCDLCLMALDOSQPKISRHLAMLRSGILLDRKQGW 60

QY 69 VHYRLSPHSPWAAQIEQAW 89
Db 61 IHRLSPHPAWAETIMTSW 81

RESULT 7
US-10-282-122A-69031
; Sequence 69031, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69031
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Pseudomonas putida
;
US-10-282-122A-69165

Query Match          55.7%; Score 282; DB 12; Length 115;
Best Local Similarity 54.3%; Pred. No. 1e-26;
Matches 51; Conservative 19; Mismatches 18; Indels 6; Gaps 1;

QY 4 LTPQLQFNLSDETLGIVLLREMGELCVCDLCLMALDOSQPKISRHLAMLRSGILLDR 63
Db 1 MKPLQLFKILGQTRLDIVLLKASGELCVCDIYTAIINLSQPKTSRHLAMLRSGILLDR 60
```

```
Db      2  ITPDVFKLSDETRATALLIASIGELCVCELMALNDOPKISRHLAQLRSNGMLLDR 61
Qy      64  KQKQWVHYRLSPHIPSNAQIIIE-----QAWLS 91
      :||:||||:||||:||||:||||:||||:||||:
Db      62  RQGWVYVYRLNPELPSVWHEMLQVTLQANSQWLA 95

RESULT 9
US-10-282-122A-66368
; Sequence 66368, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66368
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66368

Query Match      52.6%; Score 266; DB 12; Length 116;
Best Local Similarity 51.6%; Pred. No. 9.6e-25;
Matches 48; Conservative 22; Mismatches 23; Indels 0; Gaps 0;

Qy      5  TPLQLFKNLSDETRGIVLLREMGELCVCELMALNDOPKISRHLAQLRSNGMLLDR 64
      :||:||||:||||:||||:||||:||||:||||:
Db      3  SPAEVFKCIADETRVRATLIVDQGLCVCELMALNDOPKISRHLAQLRSAGILLDR 62

Qy      65  QGKQWVHYRLSPHIPSNAQIIIEQAWLSQQDDVQ 97
      :||:||||:||||:||||:||||:||||:
Db      63  QGQWVYVYRLNPAWIHEVQLVTLRANGDWLQ 95

RESULT 10
US-10-282-122A-77221
; Sequence 7721, Application US/10282122A
```

```
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77221
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77221

Query Match      45.1%; Score 228; DB 12; Length 113;
Best Local Similarity 45.0%; Pred. No. 4.5e-20;
Matches 45; Conservative 23; Mismatches 26; Indels 6; Gaps 2;

Qy      4  LTPQLFKNLSDETRGIVLLREMGELCVCELMALNDOPKISRHLAQLRSAGILLDR 63
      :||:||||:||||:||||:||||:||||:||||:
Db      1  MLPHQFFKLLADETRVRCCLMLAAREEKVCVAELTEALNESQPKISRHLAQLRSAGVVVDI 60

Qy      64  KQKQWVHYRLSPHIPSNAQ---IIEQAWLSQ--QDDVQ 97
      :||:||||:||||:||||:||||:||||:
Db      61  RQGWVYVYRLSDQLPGWVRKQIQGLVESNCLKQEVYQDDIQ 100

RESULT 11
US-10-282-122A-44793
; Sequence 44793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
```

```

; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44793
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-44793

Query Match 44.9%; Score 227; DB 12; Length 109;
Best Local Similarity 49.4%; Pred. No. 5.8e-20;
Matches 41; Conservative 18; Mismatches 24; Indels 0; Gaps 0;

Qy 4 LTPQLQFNLSDETRIGVILL--REMGELCVCDLQALDQSQPKISRHLAMLRSGILLDR 63
Db 2 INQVDFKLSDDQRLNLKLVNKNQVCLETEQLSQPKISRHLALRIHGVLLDE 61
Qy 64 KQKWHYRLSPHIPSWAAQII 86
Db 62 RKGQWVYSLNPDLPVWALDILK 84

RESULT 12
US-10-156-761-11291
; Sequence 11291, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11291
; LENGTH: 134

```

```

; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11291

Query Match 31.6%; Score 160; DB 14; Length 134;
Best Local Similarity 41.2%; Pred. No. 1.4e-11;
Matches 33; Conservative 17; Mismatches 28; Indels 2; Gaps 1;

Qy 8 QLFKNLSDETRIGVILL--REMGELCVCDLQALDQSQPKISRHLAMLRSGILLDRQK 65
Db 37 KVFKAIGDPVRLRLSLMTASRAGGEVCCDLTPDFDLQPTISHHLKLRLQAGLDCRR 96
Qy 66 GKWHYRLSPHIPSWAAQII 85
Db 97 GTWVYVLPVEMTDRLASIL 116

RESULT 13
US-10-282-122A-62327
; Sequence 62327, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62327
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-10-282-122A-62327

Query Match 29.8%; Score 151; DB 12; Length 126;
Best Local Similarity 42.6%; Pred. No. 1.6e-10;
Matches 29; Conservative 15; Mismatches 22; Indels 2; Gaps 1;

Qy 10 FKNLSDETRIGV--LLLRWELCVCDLQALDQSQPKISRHLAMLRSGILLDRQK 67

```

Db 35 FKALADPVRQLQLSSVASRAGGACVCDISAGVEVSQPTISHHLKVLRLDAGLLTERRRAS 94

QY 68 WWHYRLSP 75
||:|:|

Db 95 WVYAVVP 102

RESULT 14

US-10-282-122A-64752
; Sequence 64752, Application US/10282122A
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

Query Match	29.8%; Score 151; DB 12; Length 126;
Best Local Similarity	42.6%; Pred.No.1.6e-10;
Matches	29; Conservative 15; Mismatches 22; Indels 2; Gaps 1;
Qy	10 FKNISDETRIGIV--LLIRENGEICVCDLCAWDQSOPKISRHLAMRESGILLDRKQG 67 : : : : : : : : : : : : : :
Dd	35 FKALADPVLQLLSVASRAGGEACVDISAGVEVSQPTTISHHLKVLRDAGLLTSRRAS 94
Qy	68 WWHYRLSP 75
Dd	95 WWHYAVVP 102

RESULT 15

RESULT IS
US-10-156-761-7873
; Sequence 7873, Application US/10156761

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:34:56 ; Search time 21 Seconds
(without alignments)

444.313 Million cell updates/sec

Title: US-10-676-296-2_COPY_1_97

Perfect score: 506

Sequence: 1 MLQLTPLQLFKNLSDERLFG.....PSWAAQIIIEQAWLSQQDDVQ 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: Pirl.*

2: Pirl.*

3: Pirl.*

4: Pirl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	117	1	BVECAR
2	423	83.6	117	2	E86021
3	423	83.6	117	2	E91175
4	414	81.8	117	2	A56269
5	399	78.9	113	2	AD0406
6	266	52.6	116	2	P83361
7	228	45.1	113	2	G82245
8	181	35.8	122	2	B69220
9	176	34.8	108	2	G70420
10	166.5	32.9	108	2	AG2151
11	156.5	30.9	104	2	S74901
12	151	29.8	126	2	F70964
13	147.5	29.2	122	2	T36631
14	147.5	29.2	325	2	D87514
15	147	29.1	89	2	D64465
16	147	29.1	116	2	H84024
17	142	28.1	112	2	AD1944
18	137.5	27.2	108	2	H72306
19	136.5	27.0	111	2	F69779
20	136	26.9	111	2	AI1212
21	136	26.9	115	2	E90443
22	133	26.3	102	2	D84024
23	131	25.9	122	2	E84154
24	130.5	25.8	337	2	AD2837
25	130.5	25.8	337	2	F97634
26	129	25.5	338	2	AH3321
27	127	25.1	105	2	H69949
28	124.5	24.6	115	2	AI2776
29	122.5	24.2	129	2	T08342

30 122 24.1 122 2 D97176
31 120.5 23.8 118 2 D75141
32 120.5 23.8 135 2 E70585
33 119.5 23.6 110 2 F95983
34 119 23.5 119 2 G69106
35 118.5 23.4 115 2 C95333
36 118 23.3 125 2 A75601
37 113.5 22.4 309 2 A82658
38 112.5 22.2 122 2 S31197
39 112.5 22.2 134 2 C75193
40 111.5 22.0 104 1 B41903
41 110.5 21.8 104 1 A41902
42 110.5 21.8 104 2 F89962
43 110.5 21.8 140 2 C87012
44 109 21.5 116 2 AI2982
45 109 21.5 117 2 F98300

ALIGNMENTS

RESULT 1

BVECAR

arsenical resistance operon repressor - Escherichia coli plasmid R773

C/Species: Escherichia coli

C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999

C/Accession: JS0448

R/San Francisco, M.J.D.; Hope, C.L.; Owolabi, J.B.; Tisa, L.S.; Rosen, B.P.

Nucleic Acids Res. 18, 619-624, 1990

A/Title: Identification of the metalloregulatory element of the plasmid-encoded arsenica

A/Reference number: JS0448; MUID:90174986; PMID:2408017

A/Accession: JS0448

A/Molecule type: DNA

A/Residues: 1-117 <SAN>

A/Cross-references: GB:X16045; NID:942716; PIDN:CAA34168.1; PID:942717

C/Comment: This is a transcriptional repressor for the ars operon; it is a trans-acting

C/Genetics:

A/Gene: arsR

A/Genome: plasmid

A/Superfamily: arsenical resistance operon repressor

C/Keywords: DNA binding; homodimer; repressor; transcription regulation

Query Match	100.0%;	Score 506;	DB 1;	Length 117;	
Best Local Similarity	100.0%;	Pred. No. 1.6e-50;			
Matches	97;	Conservative	0;	Mismatches 0;	
				Indels 0;	
				Gaps 0;	
Qy	1	MLQLTPLQLFKNLSDERLGI	VLLIREMGEICVCDL	CWALDQSQPKISRHLAMLR	ESGIL 60
Db	1	MLQLTPLQLFKNLSDERLGI	VLLIREMGEICVCDL	CWALDQSQPKISRHLAMLR	ESGIL 60
Qy	61	LDRKQGWVHYRLSPHIP	SWAAQIIIEQAWLSQQDDVQ	97	
Db	61	LDRKQGWVHYRLSPHIP	SWAAQIIIEQAWLSQQDDVQ	97	

RESULT 2

E86021

hypothetical protein arsR [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C/Species: Escherichia coli

C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001

C/Accession: E86021

R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A/Reference number: A85480; MUID:21074935; PMID:11206551

A/Accession: E86021

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-117 <STO>

A/Cross-references: GB:AE005174; NID:g12518198; PIDN:AGS58633.1; GSPDB:GN00145; UWGP:Z49

A/Experimental source: strain O157:H7, substrain EDL933

C/Genetics:

R; Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
Science 277, 1453-1462, 1997
.A.; Rose, D.J.; Mau, B.; Shao, Y.
A; Title: The complete genome sequence of Escherichia coli K-12.
A; Reference number: A61720; UID:97426617; PMID:978503
A; Accession: H65147
A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Residues: 1-117 <BLAT>
A; Cross-references: GB:AE000426; GB:U00096; NID:gl789910; PIDN:AAC76526.1; PID:gl789916;
A; Experimental source: strain K-12, substrain MG1655
C; Genetics:
A; Gene: araR
A; Map position: 77.5 min
C; Superfamily: arsenical resistance operon repressor

Query Match 81.8%; Score 414; DB 2; Length 117;
Best Local Similarity 80.9%; Pred. No. 5.3e-40;
Matches 76; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY 4 LTPQLQFKMLSDTRIGIVLLIREMGELCVCDLCMALDOSQPKISRHLAMLRSGILLDR 63
| | | | : | | | : | | | | | | | | | | | | | | | | | |
Db 4 LLPIQIFKLADSTRIGIVLLISELGCVCDCITALDOSQPKISRHLALLRESGILLDR 63
| | | | : | | | : | | | | | | | | | | | | | | | | | |

QY 64 KQGKWVHYRLSPHPISWAAQIIQAWLSQQDDVQ 97
| | | | : | | | : | | | | | | | | | | | | | | | | | |
Db 64 KQGKWVHYRLSPHIPAWAAKIIDEAWRCQEKKVQ 97
| | | | : | | | : | | | | | | | | | | | | | | | | | |

RESULT 5
AD0406
arsenical resistance operon repressor [imported] - Yersinia pestis (strain CO92)
C; Species: Yersinia pestis
C; Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C; Accession: AD0406
R; Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
ill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A; Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A; Reference number: AB0001; UID:21470413; PMID:11586360
A; Accession: AD0406
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-113 <KUR>
A; Cross-references: GB:AL590842; PIDN:CAC92576.1; PID:gl5981273; GSPDB:GN00175
C; Genetics:
A; Gene: araR
C; Superfamily: arsenical resistance operon repressor

Query Match 78.9%; Score 399; DB 2; Length 113;
Best Local Similarity 76.0%; Pred. No. 2.6e-36;
Matches 73; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLQTPLQLPFKNISDETRLGVILIRMGELCVCDLCMALDOSQPKISRHLAMLRSGILL 60
| | | | : | | | : | | | | | | | | | | | | | | | | | |
Db 1 MTTLUTPLQLFNLSDTRFLMIILLKASGELCVCELCHRLNEAQPKISRHLAMLRSGILL 60
| | | | : | | | : | | | | | | | | | | | | | | | | | |

QY 61 LDRKQGWVHYRLSPHPISWAAQIIQAWLSQQDDV 96
| | | | : | | | : | | | | | | | | | | | | | | | | | |
Db 61 LDREAGKWVHYRLSPHIPAWAAAIIEQTYLSQRDEI 96
| | | | : | | | : | | | | | | | | | | | | | | | | | |

RESULT 6
F83361
ArsR protein PA2277 [imported] - Pseudomonas aeruginosa (strain PAO1)
C; Species: Pseudomonas aeruginosa
C; Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C; Accession: F83361
R; Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A;Molecule type: DNA
A;Residues: 1-122 <MTH>
A;Cross-references: GB:AE000865; GB:AE000666; NID:g2621984; PIDN:AA859397.1; PID:g2621999
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH899
C;Superfamily: arsenical resistance operon repressor

Query Match 35.8%; Score 181; DB 2; Length 122;
Best Local Similarity 51.9%; Pred. No. 2.2e-13;
Matches 40; Conservative 11; Mismatches 24; Indels 2; Gaps 2;

Qy 11 KNLSDETRLGIVLLREMGELCVCDLCMALDQSQPKISRHLAMLRSGILLDRKQKVVH 70
Db 43 KALADPTLLIYYLSE-GDLVCVEIMALKKPKOPTISHHLNILLRAGFLKAEKRGWVVH 101
Qy 71 YRL-SPHIPSWAAQIIE 86
Db 102 YSLASDDLPSMIKQVIE 118

RESULT 9
G70420
transcription regulator ArsR family - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 26-Aug-1999
C;Accession: G70420
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ovi
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: G70420
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-102 <AOF>
A;Cross-references: GB:AE000737; NID:g2983782; PIDN:AC07355.1; PID:g2983797; GB:AE000655
A;Experimental source: strain VFS
C;Genetics:
A;Gene: arsR
C;Superfamily: arsenical resistance operon repressor

Query Match 34.8%; Score 176; DB 2; Length 102;
Best Local Similarity 49.3%; Pred. No. 6.8e-13;
Matches 35; Conservative 13; Mismatches 23; Indels 0; Gaps 0;

Qy 8 QLEKNLSDETRIGIVLLRMGELCVCDLCMALDQSQPKISRHLAMLRSGILLDRKQK 67
Db 8 RIFYALSPPKRLCWKLLERELCVCDPMRIFKESQPKISFHLKVLREAGLVTQKRGK 67
Qy 68 WVVYRLSPHIP 78
Db 68 WNYVRLNKEAP 78

RESULT 10
AG2151
transcription regulator alr2766 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AG2151
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.; Iriguchi
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG2151
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-108 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA874465.1; PID:g17131859; GSPDB:GN00179
A;Experimental source: strain PCC 7120

C:Genetics: alr2766
A:Gene: RV2542
C:Superfamily: arsenical resistance operon repressor

Query Match 32.9%; Score 166.5; DB 2; Length 108;
Best Local Similarity 47.2%; Pred. No. 8.8e-12;
Matches 34; Conservative 16; Mismatches 19; Indels 3; Gaps 2;

QY 10 FKNLSDETRGLGIVLLREMGELCVCDLQMALDQSQPKISRHLAMLRRESGILLDRKQKV 69
Db 16 FHALSDPIRLQVLTLLR-ELCVCDLQMALDQSQPKISRHLAMLRRESGILLDRKQKV 74

QY 70 HYLSPSPSWA 81
Db 75 YVSL--NLPQFA 84

RESULT 11
S74901
arsenical resistance operon repressor - *Synechocystis* sp. (strain PCC 6803)
N:Alternate names: protein sll1957
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S74901
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, K.; K. 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74901
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-104 <KAN>
A:Cross-references: EMBL:D90909; GB:AB001339; NID:g1652844; PIDN:BAAL7862.1; PID:g165294
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: arsa
A:Start codon: GTG
C:Superfamily: arsenical resistance operon repressor

Query Match 30.9%; Score 156.5; DB 2; Length 104;
Best Local Similarity 48.5%; Pred. No. 1.2e-10;
Matches 32; Conservative 14; Mismatches 19; Indels 1; Gaps 1;

QY 10 FKNLSDETRGLGIVLLREMGELCVCDLQMALDQSQPKISRHLAMLRRESGILLDRKQKV 69
Db 13 FOALSDPIRLQVLTLL-ENQEQVCVCDLQDQNISQSKLSFHLKRLDAELVHTRQDGRWI 71

QY 70 HYLSP 75
Db 72 YVRLNP 77

RESULT 12
F70964
hypothetical protein RV2642 - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70964
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:9829987; PMID:9634230
A:Accession: F70964
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-126 <COL>
A:Cross-references: GB:Z80225; GB:AL123456; NID:g3242265; PIDN:CAB02348.1; PID:g1550683

A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2542
C:Superfamily: arsenical resistance operon repressor

Query Match 29.8%; Score 151; DB 2; Length 126;
Best Local Similarity 42.6%; Pred. No. 6.2e-10;
Matches 29; Conservative 15; Mismatches 22; Indels 2; Gaps 1;

QY 10 FKNLSDETRGLGIV--LLIREMGELCVCDLQMALDQSQPKISRHLAMLRRESGILLDRKQKV 67
Db 35 FXALADPVRQLLSVSRAGGACVCDISAGVEVSQPTISHHLKVLDRDAGLTSRRAS 94

QY 68 WVHYRLSP 75
Db 95 WYVAVVP 102

RESULT 13
T36631
probable transcription regulator - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T36631
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21610
A:Accession: T36631
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-122 <OLI>
A:Cross-references: EMBL:AL078610; PIDN:CAB44399.1; GSPDB:GN00070; SCOEDB:SCH35.28C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCH35.28C
C:Superfamily: arsenical resistance operon repressor

Query Match 29.2%; Score 147.5; DB 2; Length 122;
Best Local Similarity 37.5%; Pred. No. 1.5e-09;
Matches 30; Conservative 18; Mismatches 29; Indels 3; Gaps 2;

QY 8 QLFKNLSDETRGLGIV--LLIREMGELCVCDLQMALDQSQPKISRHLAMLRRESGILLDRKQ 65
Db 41 RMFKALGDFVRLRLFSVASHEGGEACVCDI-SDVGVSQPTVSHHLKVLKLEAGLTSERR 99

QY 66 GKVVHYRLSPHIPSAAQII 85
Db 100 GTWVYTRVPSVLAMGQLL 119

RESULT 14
D87514
transcription regulator, ArsR family [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: D87514
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; Laub, M.C.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <STO>
A:Cross-references: GB:AE005673; NID:g13423632; PIDN:AAK24112.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2141

Query Match 29.2%; Score 147.5; DB 2; Length 325;
Best Local Similarity 40.2%; Pred. No. 4.6e-09;
Matches 33; Conservative 17; Mismatches 29; Indels 3; Gaps 2;

Search completed: June 2, 2004, 19:38:48
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:31:11 ; Search time 17 Seconds
(without alignments)
297.106 Million cell updates/sec

Title: US-10-676-296-2_COPY_1_97
Perfect score: 506
Sequence: 1 MLQITPLQLFKNLSDERLIG.....PSWAAQIIIEQAWLSQQDDVQ 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

* SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	506	100.0	117	1 ARR1_ECOLI	P15905 escherichia
2	447	88.3	117	1 ARR2_ECOLI	P52144 escherichia
3	414	81.8	117	1 ARSR_ECOLI	P37309 escherichia
4	163.5	32.3	121	1 CADF_STRAU	P37374 staphylococ
5	147	29.1	89	1 YD25_METUA	Q58721 methanococc
6	145	28.7	119	1 CADC_LISMO	Q56405 listeria mo
7	129	25.5	119	1 CADC_LISIN	P94887 listeria in
8	127	25.1	105	1 ARSR_BACSU	P45949 bacillus su
9	112.5	22.2	122	1 SMTB_SYNP7	P30340 synechococc
10	111.5	22.0	104	1 ARSR_STAXU	P30338 staphylococ
11	110.5	21.8	104	1 ARSR_STAXY	Q01256 staphylococ
12	108.5	21.4	108	1 HUYU_VIBCH	P32695 vibrio chol
13	108.5	21.4	122	1 CADC_STRAU	P20047 staphylococ
14	107.5	21.2	122	1 CADC_BACPF	P30339 bacillus ps
15	106	20.9	132	1 SMTB_SYNP3	Q55940 synechocyst
16	82.5	16.3	125	1 MERR_STRLI	P30346 streptomyce
17	80.5	15.9	157	1 YF53_METUA	Q58948 methanococc
18	72	14.2	118	1 Y094_MYCTU	Q10864 mycobacteri
19	68.5	13.8	644	1 DXS_CHELE	Q02619 chlamydia p
20	68.5	13.5	514	1 IMDH_LEIDO	P21620 leishmania
21	68.5	13.5	1620	1 ALK_HUMAN	Q9um73 homo sapien
22	68	13.4	409	1 Y774_METUA	Q58184 methanococc
23	68	13.4	910	1 RDL2_ARYTH	Q8w3k3 arabidopsis
24	67.5	13.3	1078	1 GRB_SYNP3	P77966 synechocyst
25	66	13.0	744	1 ATKB_RALSO	Q58233 methanococc
26	64.5	12.7	257	1 Y823_METUA	P87503 human adeno
27	64	12.6	1193	1 DPOL_ADE04	Q31178 bacillus an
28	63.5	12.5	99	1 PAGR_BACAN	P50098 trypanosoma
29	63.5	12.5	512	1 IMDH_TRYEB	Q9ae11 corynebacte
30	63.5	12.5	610	1 PPCK_CORGL	Q58233 methanococc
31	63	12.5	208	1 Y379_METUA	Q7824 methanococc
32	62.5	12.4	308	1 GDFH_HUMAN	Q95988 homo sapien
33	62.5	12.4	1082	1 A3B2_HUMAN	Q13367 homo sapien

RESULT 1
ARR1_ECOLI STANDARD; PRT; 117 AA.
AC P15905;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Arsenical resistance operon repressor.
GN ARSR.
OS Escherichia coli.
OG Plasmid R773.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90174986; PubMed=2408017;
RA San Francisco M.J.D., Hope C.D., Owolabi J.B., Tisa L.S., Rosen B.P.;
RT "Identification of the metalloregulatory element of the
plasmid-encoded arsenical resistance operon."
RL Nucleic Acids Res. 18:619-624(1990).
RN [2]
RP FUNCTION.
RX MEDLINE=92157859; PubMed=1838573;
RA Wu J., Rosen B.P.;
RT "The Arsr protein is a trans-acting regulatory protein."
RL Mol. Microbiol. 5:1331-1336(1991).
RN [3]
RP METAL-REGULATION.
RX MEDLINE=93107054; PubMed=8416957;
RA Wu J., Rosen B.P.;
RT "Metalloregulated expression of the ars operon."
RL J. Biol. Chem. 268:52-58(1993).
CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR FOR THE ARS OPERON. ARSR IS
A TRANS-ACTING REGULATORY PROTEIN WHICH CONTROLS ITS OWN
EXPRESSION. THE REPRESSIVE EFFECT OF ARSR IS ALLEVIATED BY OXYGENS
OF +III OXIDATION STATE OF ARSENIC, ANTIMONY, AND BISMUTH, AS WELL
AS ARSENATE (AS(V)).
CC -!- SUBUNIT: Binds DNA as a homodimer.
CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
REGULATORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.ebi.ac.uk/ebis-sib.ch/>)
or send an email to license@sib-sib.ch.

EMBL: X16045; CAA34168.1; --
PIR: JS0448; BVECAR.
HSSP: P30340; 1SMT.
InterPro: IPR001845; HTH_ArsR.
Pfam: PF01022; HTH_5; 1.
PRINTS: PR00778; HTHARSR.

P24008 rattus norv
P56114 helicobacte
Q10283 schizosacch
P25243 bacterioph
Q8tym5 methanopyru
P33400 saccharomyc
Q08493 homo sapien
P43376 escherichia
P97793 mus musculu
Q10187 schizosacch
P33233 escherichia
Q7vn67 haemophilus

DR SMART; SM00418; HTH_ARSR; 1.
 DR PROSITE; PS00846; HTH_ARSR_FAMILY; 1.
 KW Plasmid; Arsenical resistance; Transcription regulation; Repressor;
 FT DNA BIND 33 52 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 117 AA; 13198 MW; 1F0D10766B4FD886 CRC64;
 Query Match 100.0%; Score 506; DB 1; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2e-48;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLQLTPLQLFKNLSDTRIGIVLLREMGELCVCDLCLMALDOSQPKISRHLMRESGIL 60
 Db 1 MPEIASLQFKLSDETRIGIVLLREMGELCVCDLCLCTALEQSQPKTSHLMRESGIL 60

QY 61 LDRKQKQWVHYRLSPHPSWAAQIIIEQAWLSQDDVQ 97
 Db 61 LDRKQKQWVHYRLSPHPSWAAQIIIEQAWLSQDDVQ 97

RESULT 2
 ARSR_ECOLI
 ID ARSR_ECOLI STANDARD; PRT; 117 AA.
 AC P52144;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Arsenical resistance operon repressor.
 GN ARSR.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96275894; PubMed=8674982;
 RX Bruhn D.P., Li J., Silver S., Roberto F., Rosen B.P.;
 RA "The arsenical resistance operon of IncN plasmid R46.";
 RL FEMS Microbiol. Lett. 139:149-153(1996).
 CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR FOR THE ARS OPERON. ARSR IS
 CC EXPRESSION. THE REPRESSIVE EFFECT OF ARSR IS ALLEVIATED BY OXYIONS
 CC OF +III OXIDATION STATE OF ARSENIC, ANTIMONY, AND BISMUTH, AS WELL
 CC AS ARSENATE (AS(V)).
 CC -1- SUBUNIT: Binds DNA as a homodimer.
 CC -1- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL; U38947; AAB09624.1; --
 DR HSP; P30340; ISMT.
 DR InterPro; IPR001845; HTH_Arsr.
 DR Pfam; PF01022; HTH 5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SM00418; HTH_ARSR; 1.
 DR PROSITE; PS00846; HTH_ARSR_FAMILY; 1.
 KW Plasmid; Arsenical resistance; Transcription regulation; Repressor;
 FT DNA BIND 33 52 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 117 AA; 12999 MW; 4E2D132F1F011A76 CRC64;
 Query Match 88.3%; Score 447; DB 1; Length 117;
 Best Local Similarity 87.6%; Pred. No. 5.7e-42;
 Matches 85; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MLQLTPLQLFKNLSDTRIGIVLLREMGELCVCDLCLMALDOSQPKISRHLMRESGIL 60
 Db 1 MPEIASLQFKLSDETRIGIVLLREMGELCVCDLCTALEQSQPKTSHLMRESGIL 60

QY 61 LDRKQKQWVHYRLSPHPSWAAQIIIEQAWLSQDDVQ 97
 Db 61 LDRKQKQWVHYRLSPHPSWAAQIIIEQAWLSQDDVQ 97

RESULT 3
 ARSR_ECOLI
 ID ARSR_ECOLI STANDARD; PRT; 117 AA.
 AC P37309;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Arsenical resistance operon repressor.
 GN ARSR OR ARSE OR B3501.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes.";
 RL Nucleic Acids Res. 22:2576-2586(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=95238276; PubMed=7721697;
 RA Diorio C., Cai J., Marmor J., Shinder R., Dubow M.S.;
 RT "An Escherichia coli chromosomal ars operon homolog is functional in
 RT arsenic detoxification and is conserved in Gram-negative bacteria.";
 RL J. Bacteriol. 177:2050-2056(1995).
 CC -1- FUNCTION: TRANSCRIPTIONAL REPRESSOR FOR THE ARSEFG OPERON. ARSE IS
 CC A TRANS-ACTING REGULATORY PROTEIN WHICH CONTROLS ITS OWN
 CC EXPRESSION. THE REPRESSIVE EFFECT OF ARSE IS ALLEVIATED BY OXYIONS
 CC OF +III OXIDATION STATE OF ARSENIC, ANTIMONY, AND BISMUTH, AS WELL
 CC AS ARSENATE (AS(V)) (BY SIMILARITY).
 CC -1- SUBUNIT: Binds DNA as a homodimer (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL; U00039; AAB18477.1; --
 DR EMBL; AE000426; AAC76526.1; --
 DR EMBL; X80057; CAA56361.1; --
 DR PIR; A56269; A56269.
 DR HSP; P30340; ISMT.
 DR EcoGene; EG12235; arsr.
 DR InterPro; IPR001845; HTH_Arsr.
 DR Pfam; PF01022; HTH 5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SM00418; HTH_ARSR; 1.
 DR PROSITE; PS00846; HTH_ARSR_FAMILY; 1.
 KW Arsenical resistance; Transcription regulation; Repressor;
 FT DNA BIND 33 52 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 117 AA; 13252 MW; 35BC3F6F94BBD3DB CRC64;
 Query Match 81.8%; Score 414; DB 1; Length 117;
 Best Local Similarity 80.9%; Pred. No. 2.3e-38;

Matches 76; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY 4 LTPQLFKNLSDETRGLGIVLLRMEGLCVCDLQWALDQSQPKISRHLAMLRSGILLDR 63
 Db 4 LLPTQLFKILADETRGLGIVLLRMEGLCVCDLQWALDQSQPKISRHLAMLRSGILLDR 63

QY 64 KQGRWVHRLSPHIPSWAAQIIQEWLQWLSQQDDVQ 97
 Db 64 KQGRWVHRLSPHIPSWAAQIIQEWLQWLSQQDDVQ 97

RESULT 4
 CADF STAAU STANDARD; PRT; 121 AA.

AC P37374;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cadmium efflux system accessory protein homolog.
 GN CADF.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]

SEQUENCE FROM N.A.
 TRANSPOSON=PaIt554;
 RA Chikranne S.G., Dubin D.T.;
 RA "PaIt554": A Staphylococcus aureus chromosomal element encoding
 RT cadmium resistance determinants, and genes resembling the transposases
 RT genes of Tn554";
 RL Submitted (FBI-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

 CC EMBL; L10909; AAA26609.1; -;
 DR HSP; P30340; 1SMT.
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SMO0418; HTH ARSR; 1.
 DR PROSITE; PS00846; HTH ARSR FAMILY; 1.
 KW Transcription regulation; DNA-binding; Cadmium resistance;
 FT DNA BIND 58 77 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 121 AA; 13693 MW; F6380889DB94486C CRC64;

Query Match 32.3%; Score 163.5; DB 1; Length 121;
 Best Local Similarity 40.9%; Pred. No. 5.5e-11;
 Matches 36; Conservative 19; Mismatches 28; Indels 5; Gaps 2;

QY 8 QLFKNLSDETRGLGIVLLRMEGLCVCDLQWALDQSQPKISRHLAMLRSGILLDRKQK 67
 Db 33 KVFKAISDDTRVKIAYVLSLEGELCVCDVANIIESSTATSHLLRLNKLGIKYKKGK 92

QY 68 WYHRL-SPHIPSWAAQIIQEWLQWLSQQD 94
 Db 93 LVYISLDDHEV---KQLVKAFLQRE 116

RESULT 5
 YD25 METJA STANDARD; PRT; 89 AA.
 ID YD25_METJA
 AC Q58721;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)

15-MAR-2004 (Rel. 43, Last annotation update)
 Putative HTH-type transcriptional regulator MJ1325.
 GN Methanococcus jannaschii.
 OS Methanococcus jannaschii.
 OC Methanocaldococcales; Methanococci; Methanococcales;
 OX NCBI_TaxID=2190;
 RN [1]

SEQUENCE FROM N.A.
 STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Bult C.J., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RL Science 273:1058-1073 (1996).
 CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

 CC EMBL; U67573; AAB99335.1; -;
 DR PIR; D64465; D64465.
 DR TIGR; MJ1325; -;
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SMO0418; HTH ARSR; 1.
 DR PROSITE; PS00846; HTH ARSR FAMILY; FALSE NEG.
 KW Hypothetical protein; Transcription regulation; DNA-binding;
 SQ SEQUENCE 89 AA; 10292 MW; 33AFC876DCD861F9 CRC64;

Query Match 29.1%; Score 147; DB 1; Length 89;
 Best Local Similarity 45.3%; Pred. No. 2.5e-09;
 Matches 29; Conservative 12; Mismatches 23; Indels 0; Gaps 0;

QY 8 QLFKNLSDETRGLGIVLLRMEGLCVCDLQWALDQSQPKISRHLAMLRSGILLDRKQK 67
 Db 9 EIFKAFGDPTRLMIKLLAENGSMVCVKIIDLKKPQPTISHLLNLLKAGIVKARKEGT 68

QY 68 WYH 71
 Db 69 WYF 72

RESULT 6
 CADF LISMO STANDARD; PRT; 119 AA.
 ID CADF_LISMO
 AC Q56405;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cadmium efflux system accessory protein (Cadmium resistance regulatory
 DE protein).
 GN CADF.
 OS Listeria monocytogenes.
 OG Plasmid pLM74.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=LM74;
RX MEDLINE=94245633; PubMed=8188605;
RA Lebrun M., Audurier A., Cossart P.;
RT "Plasmid-borne cadmium resistance genes in Listeria monocytogenes are
RT similar to cadA and cadC of Staphylococcus aureus and are induced by
RT cadmium.";
RL J. Bacteriol. 176:3040-3048 (1994).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=LM74; TRANSPOSON=Tn5422;
RX MEDLINE=94245634; PubMed=8188606;
RA Lebrun M., Audurier A., Cossart P.;
RT "Plasmid-borne cadmium resistance genes in Listeria monocytogenes are
RT present on tn5422, a novel transposon closely related to tn517.";
RL J. Bacteriol. 176:3049-3061 (1994).
RN [2]
CC -!- FUNCTION: Not yet known.
CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L28104; AAA25276.1; -.
CC HSSP; P30340; 1SMT.
CC InterPro; IPR001845; HTH_ArsR.
CC Pfam; PF01022; HTH_5; 1.
CC PRINTS; PR00778; HTHARSR.
CC SMART; SM00418; HTH_ARSR; 1.
CC PROSITE; PS00846; HTH_ARSR FAMILY; 1.
CC Transcription regulation; DNA-binding; Cadmium resistance; Plasmid;
CC Transposable element.
CC DNA BIND 58 77 H-T-H MOTIF (POTENTIAL).
CC SEQUENCE 119 AA; 13489 MW; 0403AC8085A4BE1 CRC64;

Query Match 28.7%; Score 145; DB 1; Length 119;
Best Local Similarity 47.8%; Pred. No. 5.7e-09;
Matches 32; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 8 QLEKNSDETRGLVLLIREMGEVCVCDLQWALDQSPKISRHLAMLRSGILLDRKQK 67
DB 33 QLEKNSDETRGLVLLIREMGEVCVCDLQWALDQSPKISRHLAMLRSGILLDRKQK 92
QY 68 WYHYLS 74
DB 93 LVYYSLA 99

RESULT 7
CADC LISIN STANDARD; PRT; 119 AA.
AC P94887;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cadmium efflux system accessory protein (Cadmium resistance regulatory
DE protein).
DE NCBI_TaxID=1642, 1360;
GN CADC OR PL10060.
OS Listeria innocua, and
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OG Plasmid pLI100, Plasmid pND302, and Plasmid pAH82.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642, 1360;
RN [1]
RN SEQUENCE FROM N.A.
RC SPECIES=L.innocua; STRAIN=CLIP 11262 / Serovar 6a; PLASMID=PL1100;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rueniok C., Amend A.,

```

```

RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Gaebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852 (2001).
RN [2]
RN SEQUENCE FROM N.A.
RC SPECIES=L.lactis; STRAIN=M71; PLASMID=pND302;
RX MEDLINE=20305032; PubMed=10844674;
RA O'Sullivan D., Twomey D.P., Coffey A., Hill C., Fitzgerald G.P.,
RA Ross P.R.;
RT "Novel type I restriction specificities through domain shuffling of
RT HsdS subunits in Lactococcus lactis.";
RL Mol. Microbiol. 36:866-875 (2000).
CC -!- FUNCTION: Not yet known.
CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL592102; CAC42058.1; -.
CC EMBL; U78967; AAB37344.1; -.
CC EMBL; AF243383; AAF98301.1; -.
CC HSSP; P30340; 1SMT.
CC InterPro; IPR001845; HTH_ArsR.
CC Pfam; PF01022; HTH_5; 1.
CC PRINTS; PR00778; HTHARSR.
CC SMART; SM00418; HTH_ARSR; 1.
CC PROSITE; PS00846; HTH_ARSR FAMILY; 1.
CC Transcription regulation; DNA-binding; Cadmium resistance; Plasmid;
CC Complete proteome.
CC DNA BIND 58 77 H-T-H MOTIF (POTENTIAL).
CC SEQUENCE 119 AA; 13382 MW; B34C01C21DC2C944 CRC64;

Query Match 25.5%; Score 129; DB 1; Length 119;
Best Local Similarity 46.8%; Pred. No. 3.2e-07;
Matches 29; Conservative 7; Mismatches 26; Indels 0; Gaps 0;

QY 10 EKNLSDETGLVLLIREMGEVCVCDLQWALDQSPKISRHLAMLRSGILLDRKQK 69
DB 35 EKNLSDETGLVLLIREMGEVCVCDLQWALDQSPKISRHLAMLRSGILLDRKQK 94
QY 70 HY 71
DB 95 YY 96

RESULT 8
ARSR BACSU STANDARD; PRT; 105 AA.
AC P45949;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

```


[4]
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RP MEDLINE=98128797; PubMed=9466913;
 RX Cook W.J., Kar S.R., Taylor K.B., Hall L.M.;
 RA "Crystal structure of the cyanobacterial metallothionein repressor
 RT SmtB: a model for metalloregulatory proteins.";
 RL J. Mol. Biol. 275:337-346(1998).
 CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR OF THE EXPRESSION OF THE
 CC SMTA GENE. THE COMPLEX OF DNA AND SMTB IS DISASSOCIATED BY
 CC ZINC IONS.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X64585; CA445872.1; -.
 DR PIR; S31197; S31197.
 DR PDB; 1SMT; 03-DEC-97.
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SM00418; HTH_ArsR; 1.
 DR PROSITE; PS00846; HTH_ArsR FAMILY; 1.
 DR Transcription regulation; Repressor; DNA-binding; 3D-structure.
 KW DNA BIND 62 81 H-T-H MOTIF (POTENTIAL).
 FT TURN 23 24
 FT HELIX 30 43
 FT HELIX 46 55
 FT TURN 56 57
 FT STRAND 60 60
 FT HELIX 62 69
 FT TURN 70 70
 FT HELIX 73 85
 FT TURN 86 87
 FT STRAND 89 94
 FT TURN 95 96
 FT STRAND 97 102
 FT HELIX 105 119
 SQ SEQUENCE 122 AA; 13544 MW; A3C98CE13552B93F CRC64;
 Query Match 22.2%; Score 112.5; DB 1; Length 122;
 Best Local Similarity 39.1%; Pred. No. 2.1e-05;
 Matches 27; Conservative 12; Mismatches 29; Indels 1; Gaps 1;
 QY 8 QLFKNLSDETRLGIVLLREMGEICVCDLCMALDQSQPKISRHLAMLRSGILLDRKQK 67
 Db 38 EFVAVLADPNRLLSLLAR-SELGVGLAQAGVSESAVSHQLSLRLNRLVSRKQGR 96
 QY 68 WYHRLSPH 76
 Db 97 HYYQLQDH 105
 RESULT 10
 ARSR STAAU
 ID ARSR STAAU STANDARD; PRT; 104 AA.
 AC P30338;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Arsenical resistance operon repressor.
 GN ARSR.
 OS Staphylococcus aureus.
 OG Plasmid p1258.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 CC NCBI_TaxID=1280;

[1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92276351; PubMed=1534328;
 RX Ji G., Silver S.;
 RA "Regulation and expression of the arsenic resistance operon from
 RT Staphylococcus aureus plasmid p1258.";
 RL J. Bacteriol. 174:3684-3694(1992).
 CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR FOR THE ARS OPERON. ARSR IS
 CC A TRANS-ACTING REGULATORY PROTEIN WHICH CONTROLS ITS OWN
 CC EXPRESSION. THE REPRESSIVE EFFECT OF ARSR IS ALLEVIATED BY OXYGENS
 CC OF +III OXIDATION STATE OF ARSENIC, ANTIMONY, AND BISMUTH, AS WELL
 CC AS ARSENATE (AS(V)).
 CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M86824; AAA25636.1; -.
 DR PIR; B41903; B41903.
 DR HSP; P30340; 1SMT.
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SM00418; HTH_ArsR; 1.
 DR PROSITE; PS00846; HTH_ArsR FAMILY; 1.
 DR Plasmid; Arsenical resistance; Transcription regulation; Repressor;
 KW DNA-binding. 32 51 H-T-H MOTIF (POTENTIAL).
 FT DNA BIND 104 AA; 11863 MW; 975664A2469CC417 CRC64;
 SQ SEQUENCE 104 AA; 11863 MW; 975664A2469CC417 CRC64;
 Query Match 22.0%; Score 111.5; DB 1; Length 104;
 Best Local Similarity 40.9%; Pred. No. 2.2e-05;
 Matches 27; Conservative 11; Mismatches 27; Indels 1; Gaps 1;
 QY 9 LFKNLSDETRLGIVLLREMGEICVCDLCMALDQSQPKISRHLAMLRSGILLDRKQK 68
 Db 9 ILKILSDSSRLLELDLL-SCGELCACLLEHFQFQPTLSHMKSLVDNELVTRKQGN 67
 QY 69 VHYRLS 74
 Db 68 HWYQLN 73
 RESULT 11
 ARSR STAXY
 ID ARSR STAXY STANDARD; PRT; 104 AA.
 AC Q01256;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Arsenical resistance operon repressor.
 GN ARSR.
 OS Staphylococcus xylosum.
 OG Plasmid pSX267.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 CC NCBI_TaxID=1288;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=DSM 20267 / Isolate C2A;
 RL MEDLINE=92276350; PubMed=1534327;
 RA Rosenstein R., Peschel A., Wieland B., Goetz F.;
 RT "Expression and regulation of the antimoneite, arsenite, and arsenate
 RT resistance operon of Staphylococcus xylosum plasmid pSX267.";
 RL J. Bacteriol. 174:3676-3683(1992).
 CC -!- FUNCTION: TRANSCRIPTIONAL REPRESSOR FOR THE ARS OPERON. ARSR IS
 CC A TRANS-ACTING REGULATORY PROTEIN WHICH CONTROLS ITS OWN
 CC EXPRESSION. THE REPRESSIVE EFFECT OF ARSR IS ALLEVIATED BY OXYGENS

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 19:34:26 ; Search time 39 seconds
(without alignments)
784.750 Million cell updates/sec

Title: US-10-676-296-2_COPY_1_97

Perfect score: 506

Sequence: 1 MLQTPQLFLKNSDRLG.....PSWAAQIEQAWLSQDDVQ 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	444	87.7	116	2	Q9KJ15 Klebsiella
2	423	83.6	117	16	Q8X5Q0 escherichia
3	411	81.2	114	16	Q7UAW0 shigella fl
4	399	78.9	113	16	Q8ZBR1 yersinia pe
5	342	67.6	106	2	Q9L335 serratia ma
6	340	67.2	117	2	P74986 yersinia en
7	339	67.0	106	2	Q8GGH7 enterobacte
8	282	55.7	115	16	Q88JD1 pseudomonas
9	270	53.4	118	2	Q9EUU5 pseudomonas
10	268	53.0	128	16	Q88LKL pseudomonas
11	266	52.6	116	16	Q91IJ7 pseudomonas
12	265	52.4	118	2	Q68020 pseudomonas
13	259	51.2	111	16	Q87KK6 vibrio para
14	247	48.8	115	16	Q8DSL8 vibrio vuln
15	242	47.8	114	16	Q8EDJ5 shewanella
16	239.5	47.3	111	16	Q8DD70 vibrio vuln

17	228	45.1	113	16	Q9KT37	Q9kt37 vibrio chol
18	182	36.0	146	17	Q8PUT5	Q8put5 methanosarc
19	181	35.8	122	17	O26985	O26985 methanobact
20	176	34.8	102	16	O67394	O67394 aquifex aeo
21	173	34.2	84	2	O50591	O50591 acidiphiliu
22	166.5	32.9	104	16	Q7U826	Q7u826 synechococc
23	166.5	32.9	108	16	Q8YTF1	Q8ytf1 anabaena sp
24	165.5	32.7	115	2	Q8RA93	Q8ra93 sinorhizobi
25	164.5	32.5	109	16	Q7V6X9	Q7v6x9 prochloroco
26	160	31.6	134	16	Q82GY9	Q82gy9 streptomyce
27	156.5	30.9	104	16	P73808	P73808 synechocyst
28	155.5	30.7	126	16	Q7V1Y1	Q7v1y1 prochloroco
29	153.5	30.3	119	16	Q9L220	Q9l220 streptomyce
30	151.5	29.9	117	16	Q9L1V5	Q9l1v5 streptomyce
31	151	29.8	126	16	P71941	P71941 mycobacteri
32	151	29.8	126	16	Q7TY65	Q7ty65 mycobacteri
33	147.5	29.2	122	16	Q9X8X8	Q9x8x8 streptomyce
34	147.5	29.2	325	16	Q9A6F3	Q9a6f3 caulobacter
35	147	29.1	89	17	Q58721	Q58721 methanococc
36	147	29.1	116	16	Q9K8K6	Q9k8k6 bacillus ha
37	143.5	28.4	128	17	Q972Q8	Q972q8 sulfolobus
38	142	28.1	112	16	Q8YXV6	Q8yxv6 anabaena sp
39	141.5	28.0	121	2	Q93GK0	Q93gk0 bacillus st
40	141.5	28.0	135	16	Q97TJ8	Q97tj8 clostridium
41	139.5	27.6	103	16	Q7VA87	Q7va87 prochloroco
42	139	27.5	100	16	Q8ETD6	Q8etd6 oceanobacil
43	137.5	27.2	98	16	Q82R21	Q82r21 streptomyce
44	137.5	27.2	108	16	Q9X093	Q9x093 thermotoga
45	136.5	27.0	111	16	P96677	P96677 bacillus su

ALIGNMENTS

RESULT 1

Q9KJ15 ID Q9KJ15 PRELIMINARY; PRT; 116 AA.
AC Q9KJ15; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE ARSR.
GN ARSR.
OS Klebsiella oxytoca.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=571;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=D12;
RA Lee S.J., Choi S.H., Park J.E., Chung M.K., Lee H.S.;
RT "Sequencing and characterization of ars operon from Klebsiella oxytoca
D12 plasmid pMH12."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AF168737; AAF89638.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SMC0418; HTH_ArsR; 1.
DR PROSITE; PS00846; HTH_ArsR_FAMILY; 1.
KW DNA-binding; Plasmid.
SQ SEQUENCE 116 AA; 12939 MW; 0A2ADE59836A986B CRC64;

Query Match 87.7%; Score 444; DB 2; Length 116;
Best Local Similarity 88.5%; Pred. No. 2.3e-43;
Matches 85; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```

QY 2 LQTLPLQLEKNSLDETRIGVILLREMGELCVCDLQMALDQSQPKISRHLMLRESGILL 61
Db 1 MKUTTLQLFKYLSDETRIGVILLREMGELCVCDLQTALEQSQPKISRHLMLRESGILL 60

QY 62 DRKQGWVHYRLSPHPSWAAQIIQEAQLWSQDDVQ 97
Db 61 DPKQGWVHYRLSPHPSWAAQVIELAWLSQDDVQ 96

RESULT 2
Q8X5Q0 PRELIMINARY; PRT; 117 AA.
AC Q8X5Q0;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Transcriptional repressor of chromosomal ars operon.
GN ARSR OR Z4903 OR EGS4373.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533 (2001).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22 (2001).
DR EMBL; AE005575; AG58633.1; -
DR EMBL; AP002565; BAB37796.1; -
DR PIR; E91175; E91175.
DR PIR; E91175; E91175.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH 5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH ARSR.
DR PROSITE; PS00846; HTH_ARSR_FAMILY; 1.
KW Complete proteome.
SQ SEQUENCE 117 AA; 13340 MW; 611C9E8723EBDF66 CRC64;

Query Match 83.6%; Score 423; DB 16; Length 117;
Best Local Similarity 84.0%; Pred. No. 6.1e-41;
Matches 79; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 4 LTPLOLFKNLSDETRIGVILLREMGELCVCDLQMALDQSQPKISRHLMLRESGILLDR 63
Db 4 LTPLOLFKNLSDETRIGVILLREMGELCVCDLQTALEQSQPKISRHLMLRESGILLDR 63

QY 64 KQGWVHYRLSPHPSWAAQIIQEAQLWSQDDVQ 97
Db 64 KQGWVHYRLSPHPSWAAQIIQEAQLWSQDDVQ 97

RESULT 3
Q7UAW0 PRELIMINARY; PRT; 114 AA.
AC Q7UAW0;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Transcriptional repressor of chromosomal ars operon.
GN ARSR OR S4232.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RC MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786 (2003).
DR EMBL; AG016992; AAP19194.1; -
SQ SEQUENCE 114 AA; 12919 MW; 93EF861BC9C239D1 CRC64;

Query Match 81.2%; Score 411; DB 16; Length 114;
Best Local Similarity 78.7%; Pred. No. 1.4e-39;
Matches 74; Conservative 13; Mismatches 7; Indels 0; Gaps 0;

QY 4 LTPLOLFKNLSDETRIGVILLREMGELCVCDLQMALDQSQPKISRHLMLRESGILLDR 63
Db 1 MPTQIFKILADETRIGVILLREMGELCVCDLQTALEQSQPKISRHLMLRESGILLDR 60

QY 64 KQGWVHYRLSPHPSWAAQIIQEAQLWSQDDVQ 97
Db 61 KQGWVHYRLSPHPSWAAQIIQEAQLWSQDDVQ 94

RESULT 4
Q8ZBR1 PRELIMINARY; PRT; 113 AA.
AC Q8ZBR1;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Arsenical resistance operon repressor (Transcriptional repressor of
DE chromosomal ars operon).
GN ARSR OR YPO1346 OR Y0845.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RC MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holroyd S., Jagej K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527 (2001).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RC MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,

```

RA Featherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of *Yersinia pestis* KIM.";
 RL J. Bacteriol. 184:4601-4611 (2002).
 DR EMBL; AJ414156; CAC92576.1; -.
 DR EMBL; AE013687; AAM84430.1; -.
 DR PIR; AD0406; AD0406.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005489; F:electron transporter activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006118; F:electron transport; IEA.
 DR GO; GO:0006355; P:regulation of transcription; IEA.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR PROSITE; PS00190; CYTOCHROME_C; 1.
 DR Complete proteome.
 KW SEQUENCE 113 AA; 12821 MW; A002AFIC1B9379FB1 CRC64;

Query Match 78.9%; Score 399; DB 16; Length 113;
 Best Local Similarity 76.0%; Pred. No. 3.5e-38;
 Matches 73; Conservative 13; Mismatches 10; Indels 0; Gaps 0;
 QY 1 MQLTFLQFNLSDETRIGVILLREMGELCVCDLCAALDQSQPKISRHLAMLRSGILL 60
 DB 1 MTTLTFLQFNLSDETRINILLKASGELCVCELCHRLNEAQPKISRHLAMLRSGILL 60
 QY 61 LDRQKQKWHYRLSPHPSWAAQIIIEQAWLSQDDVQ 96
 DB 61 LDRRAGKWHYRLSPHPSWAAQIIIEQAWLSQDDVQ 96

RESULT 5
 Q9L335 PRELIMINARY; PRT; 106 AA.
 AC Q9L335;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE ArsR regulatory protein.
 GN ARSR.
 OS Serratia marcescens.
 OG Plasmid R478.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Serratia.
 OX NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ryan D.J.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ryan D.J.;
 RL Thesis (1999), Environmental Research Unit, Microbiology Department, .
 CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL; AJ288983; CAB88406.1; -.
 DR HSP; P30340; 1SMT.
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription; IEA.
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SM00418; HTH_ArsR; 1.
 DR PROSITE; PS00846; HTH_ArsR_FAMILY; 1.
 KW DNA-binding; Plasmid.
 SQ SEQUENCE 106 AA; 11997 MW; EB4F0DE141AE9C CRC64;
 Query Match 67.6%; Score 342; DB 2; Length 106;
 Best Local Similarity 62.8%; Pred. No. 1.2e-31;

Matches 59; Conservative 22; Mismatches 13; Indels 0; Gaps 0;
 QY 4 LTPQLFNLSDETRIGVILLREMGELCVCDLCAALDQSQPKISRHLAMLRSGILLDR 63
 DB 2 LQPVQLFKLSDETRIVMLRSEGMCMVCDICAAATQSQPKISRHWALLREAEIVDR 61
 QY 64 KQKQKWHYRLSPHPSWAAQIIIEQAWLSQDDVQ 97
 DB 62 REGKWHYRLSPHPSWAAQIIIDTAWNCERENIR 95
 RESULT 6
 P74986 PRELIMINARY; PRT; 117 AA.
 AC P74986;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Arsenite inducible repressor.
 GN ARSR.
 OS *Yersinia enterocolitica*.
 OG Plasmid pYve27.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Yersinia*.
 OX NCBI_TaxID=630;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=439-80; TRANSPOSON=TN2502;
 RX MEDLINE=97158651; PubMed=9006011;
 RA Neyt C., Iriarte M., Thi V.H., Cornelis G.R.;
 RT "Virulence and arsenic resistance in *Yersinia*.";
 RL J. Bacteriol. 179:612-619 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W22703;
 RA Iriarte M., Lambermont I., Kerbouch C., Cornelis G.R.;
 RT "Detailed genetic map of the pYve227 plasmid of *Yersinia*
 RT enterocolitica serotype O:9".
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL; U58366; AAB42205.1; -.
 DR EMBL; AF102990; AAD16860.1; -.
 DR HSP; P30340; 1SMT.
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription; IEA.
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SM00418; HTH_ArsR; 1.
 DR PROSITE; PS00846; HTH_ArsR_FAMILY; 1.
 KW DNA-binding; Plasmid.
 SQ SEQUENCE 117 AA; 13559 MW; CFB14B3BD3E47891 CRC64;
 Query Match 67.2%; Score 340; DB 2; Length 117;
 Best Local Similarity 69.8%; Pred. No. 2.3e-31;
 Matches 60; Conservative 15; Mismatches 11; Indels 0; Gaps 0;
 QY 4 LTPQLFNLSDETRIGVILLREMGELCVCDLCAALDQSQPKISRHLAMLRSGILLDR 63
 DB 2 LQPVQLFKLSDETRIAVLLRSEGMCMVCDICGATSESQPKISRHWALLREAEIVDR 61
 QY 64 KQKQKWHYRLSPHPSWAAQIIIEQAW 89
 DB 62 REGKWHYRLSPHPSWAAAEITTTW 87
 RESULT 7
 Q8GGH7 PRELIMINARY; PRT; 106 AA.
 AC Q8GGH7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative arsenic-efflux pump regulatory protein.
 GN ARSR.
 OS Enterobacter cloacae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Enterobacter.
 OX NCBI_TaxID=550;
 RP [1]
 RA Das S., Lehn N., Arnold M., Linde H.-J.;
 RT "Ars operon of Enterobacter cloacae.";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF521304; AA016021.1;
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SM00418; HTH_ArsR; 1.
 SQ SEQUENCE 106 AA; 12188 MW; 3B96D93A0B39A2DF CRC64;
 Query Match 67.0%; Score 339; DB 2; Length 106;
 Best Local Similarity 65.1%; Pred. No. 2.7e-31;
 Matches 56; Conservative 21; Mismatches 9; Indels 0; Gaps 0;
 YQ 4 LTPQLQFKNLSDTRILGIVLLRMEGELCVCDLCMALDQSQPKISRHLAMLRSGILLDR 63
 Db 2 LHPQLFKLSDTRILGIVLLRMEGELCVCDLCMALDQSQPKISRHLAMLRSGILLDR 61
 YQ 64 KQKQVHYRLSPHPSWAAQIIIE 89
 Db 62 REGKWIYRLSPNPAWAATVIDTSW 87
 RESULT 8
 Q88JDI PRELIMINARY; PRT; 115 AA.
 AC Q88JDI;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Arsenic resistance transcriptional regulator.
 GN ARSR-2 OR P2718.
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RP [1]
 RA Medline=22423060; PubMed=12534463;
 RX Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson M., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hobeisel J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808 (2002).
 DR EMBL; AS016784; AA068326.1;
 DR TIGR; P2718;
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 KW Complete proteome.

SQ SEQUENCE 115 AA; 13113 MW; B66F9603771FC6C6 CRC64;
 Query Match 55.7%; Score 282; DB 16; Length 115;
 Best Local Similarity 54.3%; Pred. No. 1.1e-24;
 Matches 51; Conservative 19; Mismatches 18; Indels 6; Gaps 1;
 YQ 4 LTPQLQFKNLSDTRILGIVLLRMEGELCVCDLCMALDQSQPKISRHLAMLRSGILLDR 63
 Db 2 ITPPDVFKLSDETRATLTLLASIGELCVCELMLCALNDQSQPKISRHLAQLRSGMLLDR 61
 YQ 64 KQKQVHYRLSPHPSWAAQIIIE-----QAWLS 91
 Db 62 RQGWYVYRLNPELPSWVHEMLQVTLQANSQWLA 95
 RESULT 9
 Q9EUUS PRELIMINARY; PRT; 118 AA.
 AC Q9EUUS;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Arsenic resistance operon regulator.
 GN ARSR.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TB3;
 RA de Ferra F., Pedrazzoli E.;
 RT "Optimization of the performance of chloroamatic degrading strains
 RT : construction of heavy metal resistance cassettes for Pseudomonas
 RT strains.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL; AJ271973; CAC18652.1;
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001845; HTH_ArsR.
 DR Pfam; PF01022; HTH_5; 1.
 DR PRINTS; PR00778; HTHARSR.
 DR SMART; SM00418; HTH_ArsR; 1.
 KW DNA-binding.
 SQ SEQUENCE 118 AA; 13490 MW; 87CD104A6316ACB CRC64;
 Query Match 53.4%; Score 270; DB 2; Length 118;
 Best Local Similarity 51.0%; Pred. No. 2.7e-23;
 Matches 51; Conservative 17; Mismatches 24; Indels 8; Gaps 2;
 YQ 4 LTPQLQFKNLSDTRILGIVLLRMEGELCVCDLCMALDQSQPKISRHLAMLRSGILLDR 63
 Db 5 ITPPTLFKLCADATRLTLTLREGELCVCELIALDQSQPKISRHLAQLRSGILLDR 64
 YQ 64 KQKQVHYRLSPHPSWAAQII-----EQAWLSQDDVQ 97
 Db 65 RQGWYVYRLNPELPSWVHEMLQVTLQANSQWLA 102
 RESULT 10
 Q88LKI PRELIMINARY; PRT; 128 AA.
 AC Q88LKI;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Arsenic resistance transcriptional regulator.
 GN ARSR-1 OR P21930.
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;


```

OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Sjepeandic D., Hohnesl J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duertshoef A., Thummler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016781; AAN67547.1; -.
DR TIGR; PP1930; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
KW Complete proteome.
SQ SEQUENCE 128 AA; 14487 MW; 29A5C5AA11532BC9 CRC64;

Query Match 53.0%; Score 268; DB 16; Length 128;
Best Local Similarity 54.3%; Pred. No. 5e-23;
Matches 51; Conservative 17; Mismatches 20; Indels 6; Gaps 1;

QY 4 LTPLQFKNLSDTELRGIVLLREMGELCVCDLQWALDQSQPKISRHLAMRESGILLDR 63
DB 15 LTPLVFKCLADTRATWLLIAREGELCVCELTHALQSLQPKISRHLAUREAGILLDR 74

QY 64 QGKWHYRLSPHIPSWAQAIIIE-----QAWLS 91
DB 75 RKQGVYVYRLHPEVPEQWVDMKGVGDANQEWLS 108

+RESULT 11
Q91ILJ7
ID Q91ILJ7 PRELIMINARY; PRT; 116 AA.
AC Q91ILJ7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ArsR protein.
GN ARSR OR PA2277.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AE004653; AAG05665.1; -.
DR PIR; F83361; F83361.
DR HSSP; P30340; ISMT.
GO; GO:0005622; C:intracellular; IEA.

```

```

DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
KW DNA-binding; Transcription regulation; Complete proteome.
SQ SEQUENCE 116 AA; 13045 MW; 3DDD19DC18F7091E CRC64;

Query Match 52.6%; Score 266; DB 16; Length 116;
Best Local Similarity 51.6%; Pred. No. 7.7e-23;
Matches 48; Conservative 22; Mismatches 23; Indels 0; Gaps 0;

QY 5 TPLQFKNLSDTELRGIVLLREMGELCVCDLQWALDQSQPKISRHLAMRESGILLDRK 64
DB 3 SPAEVFKCLADETRVATLLVDQGLCVCELQALADSQPKISRHLAQLRSAGILLDR 62

QY 65 QGKWHYRLSPHIPSWAQAIIIEQAWLSQQDDVQ 97
DB 63 QGQWVYVYRLNPAWITHEVLQVTLRANGDWLQ 95

+RESULT 12
O68020
ID O68020 PRELIMINARY; PRT; 118 AA.
AC O68020;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ArsR.
GN ARSR.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PAO1;
RX MEDLINE=99018820; PubMed=9802012;
RA Cai J., Salmon K., DuBow M.S.;
RT "A chromosomal ars operon homologue of Pseudomonas aeruginosa confers
RT increased resistance to arsenic and antimony in Escherichia coli.";
RL Microbiology 144:2705-2713(1998).
CC -1- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AF010234; AAC69642.1; -.
DR HSSP; P30340; ISMT.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH_ArsR; 1.
KW DNA-binding.
SQ SEQUENCE 118 AA; 13049 MW; 23EDAF846A58D52B CRC64;

Query Match 52.4%; Score 265; DB 2; Length 118;
Best Local Similarity 56.1%; Pred. No. 1e-22;
Matches 46; Conservative 20; Mismatches 16; Indels 0; Gaps 0;

QY 5 TPLQFKNLSDTELRGIVLLREMGELCVCDLQWALDQSQPKISRHLAMRESGILLDRK 64
DB 3 SPAEVFKCLADETRVATLLVDQGLCVCELQALADSQPKISRHLAQLRSAGILLDR 62

QY 65 QGKWHYRLSPHIPSWAQAIIIE 86
DB 63 QGQWVYVYRLNPAWITHEVLQ 84

+RESULT 13
Q87KK6
ID Q87KK6 PRELIMINARY; PRT; 111 AA.

```


Thu Jun 3 10:25:26 2004

us-10-676-296-2_copy_1_97.rapt

Page 7

Search completed: June 2, 2004, 19:38:15
Job time : 41 secs